


```

? RESULT 5
? US-09-743-247A-59
? Sequence 59, Application US/09743247A
? GENERAL INFORMATION:
? APPLICANT: Sagami Chemical Research Center, Protegene Inc.
? TITLE OF INVENTION: Human Proteins Having Hydrophobic Domains And DNAs Encoding The
? FILE REFERENCE: 1997.13300
? CURRENT APPLICATION NUMBER: US/09/743,247A
? PRIOR FILING DATE: 1999-07-22
? PRIOR APPLICATION NUMBER: JP 10-208820
? PRIOR FILING DATE: 1998-07-24
? PRIOR APPLICATION NUMBER: JP 10-224105
? PRIOR FILING DATE: 1998-08-07
? PRIOR APPLICATION NUMBER: JP 10-238116
? PRIOR FILING DATE: 1998-08-25
? PRIOR APPLICATION NUMBER: JP 10-254736
? PRIOR FILING DATE: 1998-09-09
? PRIOR APPLICATION NUMBER: JP 10-275505
? PRIOR FILING DATE: 1998-09-29
? NUMBER OF SEQ. ID NOS: 150
? SOFTWARE: Windows 95 (word 98)
? SEQ ID NO 59

```



```
LENGTH: 771
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (729)..(729)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (756)..(756)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-3128
```

```
Query Match
Best Local Similarity 54.1%; Score 54.4; DB 5; Length 771;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;
```

```
CY 337 GAATTCACCCGAGAGCGACTTAAGCCATTAACAGCAGCCGACGAAATCAAGCCGATCTAC 396
|||
DB 304 GACTTCACCCCGAGAGCGACTTCGCGGCGCTTGACGCGGCTCAGAGACCC--GGGCACTACTC 360
CY 397 GTCCCAATCAAGGCGCGCTGTGTTGAGATCACCACCGGAAAAATCCTTTACGCGCTCCGGA 456
|||
DB 361 ATGCCCATCAAGGCGAGGTTTGATGATGACCAAGGCGGAAATTTCTACGGGCGCCGAG 420
CY 457 GCGCATTAATCTGATGTTGCGCGGAAAGACGCGAGAGAGCTTTGGGTAAAGATGAATAG 516
|||
DB 421 GGGCGGATGCGGCTTTGCTGGAAGAGATGATCCAGGCGGCTTGCACATTTGGCTG 480
CY 517 AAGCAAGAA-----GATGTGTCCTCTCTTGAAGGTCTACGAGAAAGATC 567
|||
DB 481 GATAGGAGAGACTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
CY 568 AATACCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
|||
DB 541 GAGACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 594
```

```
RESULT 12
US-09-724-676A-3128
Sequence 3128, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3128
LENGTH: 771
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (729)..(729)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (756)..(756)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-3128
```

```
Query Match
Best Local Similarity 54.1%; Score 54.4; DB 5; Length 771;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;
```

```
CY 337 GAATTCACCCGAGAGCGACTTAAGCCATTAACAGCAGCCGACGAAATCAAGCCGATCTAC 396
|||
DB 304 GACTTCACCCCGAGAGCGACTTCGCGGCGCTTGACGCGGCTCAGAGACCC--GGGCACTACTC 360
CY 397 GTCCCAATCAAGGCGCGCTGTGTTGAGATCACCACCGGAAAAATCCTTTACGCGCTCCGGA 456
|||
DB 361 ATGCCCATCAAGGCGAGGTTTGATGATGACCAAGGCGGAAATTTCTACGGGCGCCGAG 420
CY 457 GCGCATTAATCTGATGTTGCGCGGAAAGACGCGAGAGAGCTTTGGGTAAAGATGAATAG 516
|||
DB 421 GGGCGGATGCGGCTTTGCTGGAAGAGATGATCCAGGCGGCTTGCACATTTGGCTG 480
CY 517 AAGCAAGAA-----GATGTGTCCTCTCTTGAAGGTCTACGAGAAAGATC 567
|||
DB 481 GATAGGAGAGACTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
CY 568 AATACCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
|||
DB 541 GAGACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 594
```

```
DB 361 ATGCCCATCAAGGCGAGGTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
CY 457 GCGCATTAATCTGATGTTGCGCGGAAAGACGCGAGAGAGCTTTGGGTAAAGATGAATAG 516
|||
DB 421 GAGCGTATGCGGCTTTGCTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
CY 517 AAGCAAGAA-----GATGTGTCCTCTCTTGAAGGTCTACGAGAAAGATC 567
|||
DB 481 GATAGGAGAGACTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
CY 568 AATACCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
|||
DB 541 GAGACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 594
```

```
RESULT 13
US-09-531-113-35326
Sequence 35326, Application US/09531113
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: Heck, Gregory R.
APPLICANT: La Rosa, Thomas J.
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(15761)B
CURRENT APPLICATION NUMBER: US/09/531,113
CURRENT FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 48629
SEQ ID NO 35326
LENGTH: 241
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 70039003H1
US-09-531-113-35326
```

```
Query Match
Best Local Similarity 52.4%; Score 40; DB 5; Length 241;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
```

```
CY 57 GTGCGATGAGATCGAGCGCTAGAGGATTAAGAACCAATAGGCTTTGCGTGG 156
|||
DB 52 GTGCGACACACTGTGAGAGGTGAGAGGATTAAGAGATTAAGGATTAAGGATTAAGGATTAAG 111
CY 157 AACTACATACCTCGGCTGTTAATCAATCTCGGAACAAGTATGATGATGATGATGATGATGATG 216
|||
DB 112 AACACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 171
CY 217 GGGAAAGGTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 264
|||
DB 172 GCGAAGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 219
```

```
RESULT 14
US-10-240-453-156
Sequence 159, Application US/10240453
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIERREBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
TITLE OF INVENTION: Transcription
FILE REFERENCE: 5013.1009
CURRENT APPLICATION NUMBER: US/10/240,453
CURRENT FILING DATE: 2002-10-02
PRIORITY APPLICATION NUMBER: PCT/EP01/03973
PRIORITY FILING DATE: 2001-04-06
PRIORITY APPLICATION NUMBER: DE 10019056.8
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: DE 10019171.8
PRIORITY FILING DATE: 2000-04-07
PRIORITY APPLICATION NUMBER: DE 10032529.7
```

PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: DE 10043826.1
 PRIOR FILING DATE: 2000-09-01
 NUMBER OF SEQ ID NOS: 350
 SEQ ID NO 159

LENGTH: 6801

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 10-240-453-159

Query Match: 4.9%; Score 38.4; DB 6; Length 6601;
 Best Local Similarity 58.9%; Pred. No. 1.2;
 Matches 64; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

660 TGTATGTAACATATGTCGTGAGGATCTTGTGTGTTGCTGATTCGTTG 719
 |||||
 2252 TGTATGTAACAAAAGTGTGTGTGTTGTTGGTTGGTTTGTTTTAAAGGATTT 2311

720 GATCGATCGTTTGATACATTACCAATAGTACCAATTATCTATGAAATA 771
 |||||
 2312 GTTTTGTTTTAAATAATTAATATATATAGTATATCGTTAAATA 2363

SEQ ID NO 159
 Length 198
 Sequence 198; Application US/10240485
 GENERAL INFORMATION:

APPLICANT: OLEK, Alexander

APPLICANT: PIEPERROCK, Christian

APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with

TITLE OF INVENTION: Metastasis

FILE REFERENCE: 5013.1007

CURRENT APPLICATION NUMBER: US/10/240.485

CURRENT FILING DATE: 2002-10-02

PRIOR APPLICATION NUMBER: PCT/EP01/03970

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: DE 10019058.8

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: DE 10019173.8

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 202

SEQ ID NO 198

LENGTH: 6826

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 10-240-485-198

Query Match: 4.8%; Score 38; DB 6; Length 6826;
 Best Local Similarity 57.6%; Pred. No. 1.4;
 Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

624 TGTTCCTTAGTCTCTCTTGAGATGCACTATGATGTAATGCTGTGAG 683
 |||||
 6161 TTTTAAATTTATTTGTTTGTTTTATTTTATTTTAAAGTATTTTGTTTT 6220

684 GATCTTGAGTGTGTTTCTGATTCGTTGGATCTGATCGTTTGAACAA 741
 |||||
 6221 TTTTGT 6278

Arch completed: January 8, 2003, 14:17:58
 Time: 203 secs

22 402 2 51.0 489 22 US-09-565-309A-5610 Sequence 5610, AP
23 401 2 50.9 474 22 US-09-565-309A-53681 Sequence 53681, AP
24 390 2 49.5 448 22 US-09-565-309A-5607 Sequence 5607, AP
25 366 5 46.5 367 22 US-09-565-309A-1450 Sequence 1450, AP
26 269 6 38.0 300 22 US-09-565-309A-67871 Sequence 67871, AP
27 282 6 35.8 417 22 US-09-565-309A-1375 Sequence 1375, AP
28 282 6 35.8 417 19 US-09-513-966A-8419 Sequence 8419, AP
29 282 6 35.8 477 22 US-09-565-309A-1377 Sequence 1377, AP
30 282 6 35.8 477 22 US-09-565-309A-48124 Sequence 48124, AP
31 282 6 35.8 477 22 US-09-565-309A-60955 Sequence 60955, AP
32 282 6 35.8 493 22 US-09-565-309A-1376 Sequence 1376, AP
33 282 6 35.8 8392 21 US-09-803-735-1366 Sequence 1366, AP
34 282 6 35.8 9986 20 US-09-534-555-562 Sequence 562, AP
35 282 6 35.8 9986 31 US-09-803-717-562 Sequence 562, AP
36 282 6 35.8 493 22 US-09-565-309A-1378 Sequence 1378, AP
37 276 8 35.1 413 19 US-09-513-966A-1378 Sequence 1378, AP
38 276 8 35.1 413 22 US-09-565-309A-46983 Sequence 46983, AP
39 276 8 35.1 413 24 US-09-670-3318-3290 Sequence 3290, AP
40 276 8 35.1 417 22 US-09-565-309A-42300 Sequence 42300, AP
41 272 8 34.6 420 25 US-09-649-165A-7293 Sequence 7293, AP
42 213 8 27.1 270 22 US-09-565-309A-64086 Sequence 64086, AP
43 195 6 24.8 485 25 US-09-654-617-9393 Sequence 9393, AP
44 195 6 24.8 485 27 US-09-684-016-9393 Sequence 9393, AP
45 194 6 24.7 419 17 US-09-304-517A-257238 Sequence 257238, AP

ALIGNMENTS

Seq: 1
3-03 513-996A-70554
Sequence 70554, Application US/09513996A
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING PEPTIDES
CLASS OF INVENTION: ENCODED THEREBY
P. B. REFERENCE: 2750-703P
CURRENT APPLICATION NUMBER: US/09/513,996A
CURRENT FILING DATE: 2000-02-25
NUMBER OF SEQ. ID NOS: 81028
SEQ. ID NO 70554

1. NCBI: 789
2. FE: DNA
3. ORGANISM: Arabidopsis thaliana
4. FEATURE:
5. NAME/KEY: UNSURE
6. LOCATION: 1..789
7. OTHER INFORMATION: any n or xaa = unknown
8. OTHER INFORMATION: Location 1..789 / Ceres Seq. ID 2218230
9-03 513-996A-70554

Query Match 99.9% Score 788.6; DB 19; Length 789;
Best Local Similarity 100.0%; Pred. No. 2.3e-199;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATATCAACAAACAAATTCATACACAAATTTAAACACAAAGATTTATATCTC 60
1 ATATCAACAAACAAATTCATACACAAATTTAAACACAAAGATTTATATCTC 60
61 TGAAGAAAGATGAGTCTACAGCAAGATGACAGTGGCAGTGGAGCTTGGAGCGGTA 120
61 TGAAGAAAGATGAGTCTACAGCAAGATGACAGTGGCAGTGGAGCTTGGAGCGGTA 120
121 GAGGATTTAAAGACCACTAGGCTTTTGTGGTGAACACTACATCTCCGCTTAT 180
121 GAGGATTTAAAGACCACTAGGCTTTTGTGGTGAACACTACATCTCCGCTTAT 180
181 CAATATTCGGGAACAGCTAGATCTTTTCTCAAGGAAAGAGTTCTTGTCTTCT 240
181 CAATATTCGGGAACAGCTAGATCTTTTCTCAAGGAAAGAGTTCTTGTCTTCT 240
241 GTCTCCGACCGGTACTCTCTGTGTGAGAGCAAGAAAGAGAGATTTTCTCTT 300

Db 241 GTCTCCGACCGGTACTCTCTGTGTGAGAGCAAGAAAGAGAGATTTTCTCTT 300
Cy 301 GAGAAACAAATTTGATGAGCTTTTAAAGAAAGATGATTCACAGAGCACTAAC 360
Db 301 GAGAAACAAATTTGATGAGCTTTTAAAGAAAGATGATTCACAGAGCACTAAC 360
Cy 361 CAATACACGAGCAGCAAGATCAAGAGCGATCTAGCTGCATCAAGCCGTGTTTC 420
Db 361 CAATACACGAGCAGCAAGATCAAGAGCGATCTAGCTGCATCAAGCCGTGTTTC 420
Cy 421 GAGTACCAACCGGAAATCTTTCTACGCGCTCCGAGCGCATTCATGATGTCGCGA 480
Db 421 GAGTACCAACCGGAAATCTTTCTACGCGCTCCGAGCGCATTCATGATGTCGCGA 480
Cy 481 AAAGACCGCGAGAGCTTTGGTAAAGTGAAGTGAAGACGAAGATGTTCTCTCT 540
Db 481 AAAGACCGCGAGAGCTTTGGTAAAGTGAAGTGAAGACGAAGATGTTCTCTCT 540
Cy 541 CTGAAAGTCTCACTGAGAAAGATCAATCTTTAATGATTTGGAGACCAATTTGAA 600
Db 541 CTGAAAGTCTCACTGAGAAAGATCAATCTTTAATGATTTGGAGACCAATTTGAA 600
Cy 601 GCTAAGATCTGTGTGGCGCTTGTCTCTAGAGTCTCTTGTGATGATGACTAT 660
Db 601 GCTAAGATCTGTGTGGCGCTTGTCTCTAGAGTCTCTTGTGATGATGACTAT 660
Cy 661 GTATGATCTATTTGTGTGAGATCTTTGTGTGTGTTTGTGATTTGTTGG 720
Db 661 GTATGATCTATTTGTGTGAGATCTTTGTGTGTGTTTGTGATTTGTTGG 720
Cy 721 ATCTGATCTTTTGAATACATTTACATTTACATTTACATTTACATTTACATTT 780
Db 721 ATCTGATCTTTTGAATACATTTACATTTACATTTACATTTACATTTACATTT 780
Cy 781 TTTCTGTT 789
Db 781 TTTCTGTT 789

RESULT 2
US-09-565-309A-56330
Sequence 56330, Application US/09565309A
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING PEPTIDES
CLASS OF INVENTION: THEREBY
P. B. REFERENCE: 2750-0853P
CURRENT APPLICATION NUMBER: US/09/565,309A
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ. ID NOS: 68449
SEQ. ID NO 56330
1. NCBI: 789
2. FE: DNA
3. ORGANISM: Arabidopsis thaliana
4. FEATURE:
5. NAME/KEY: misc feature
6. LOCATION: (1)..(789)
7. OTHER INFORMATION: any n = a, g, c, t, unknown, or other
8. OTHER INFORMATION: Location 1916 : OVERLAP (Clone Number : OVERLAP)
9-03 565-309A-56330

Query Match 99.9% Score 788.6; DB 22; Length 789;
Best Local Similarity 99.9%; Pred. No. 2.3e-199;
Matches 789; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ATATCAACAAACAAATTTCAATACACAAACAAACAAACAAAGATTTATATCTC 60
Db 1 ATATCAACAAACAAATTTCAATACACAAACAAACAAACAAAGATTTATATCTC 60

```

QY 61 TGAAGAGATGAGTTCTACACGAAAGATGACAGTGGCAGTGGAGATCAGACCGTA 120
DB 61 TGAAGAGATGAGTTCTACACGAAAGATGACAGTGGCAGTGGAGATCAGACCGTA 120
QY 121 GAGGCAATTAAGACCACTAGTCTTTGTGCGTGAATCTACTACTCCGTCGCTTAAT 150
DB 121 GAGGCAATTAAGACCACTAGTCTTTGTGCGTGAATCTACTACTCCGTCGCTTAAT 150
QY 181 CAACATCTCCGGAACAAGTTAGATCTGTTCTCAAGGAGAAAGTTCTCTCTCTCT 240
DB 181 CAACATCTCCGGAACAAGTTAGATCTGTTCTCAAGGAGAAAGTTCTCTCTCTCT 240
QY 241 GTCTCCGAGCCGTTACTCTCTCTGCTGAGAGCGAAGACGAAAGAACTTTTCCT 300
DB 241 GTCTCCGAGCCGTTACTCTCTCTGCTGAGAGCGAAGACGAAAGAACTTTTCCT 300
QY 301 GAGAAACATTTGATCAGAGCTTTTAAAGAAAGATGGAATTCACCGCAGAGCGTTAAG 360
DB 301 GAGAAACATTTGATCAGAGCTTTTAAAGAAAGATGGAATTCACCGCAGAGCGTTAAG 360
QY 361 CAATACCAACGACCGAGCAATCAAGCCGATCTAGTCCCAATCAAGCCGCTGTCTC 420
DB 361 CAATACCAACGACCGAGCAATCAAGCCGATCTAGTCCCAATCAAGCCGCTGTCTC 420
QY 421 GAYGTCCACACCGGAAATCCTTCTACAGGCTCCGAGGCGATTAATGATGTTCCCGGA 480
DB 421 GAYGTCCACACCGGAAATCCTTCTACAGGCTCCGAGGCGATTAATGATGTTCCCGGA 480
QY 481 AAAGACGACGACGAGCTTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
DB 481 AAAGACGACGACGAGCTTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
QY 541 CTGGAAGCTCTCAGTGAAGAAAGATCAATCTTAATGATTTGGAGACCAATTTGA 600
DB 541 CTGGAAGCTCTCAGTGAAGAAAGATCAATCTTAATGATTTGGAGACCAATTTGA 600
QY 601 GCTAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 GCTAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 TTTGCTGTT 789
DB 781 TTTGCTGTT 789

```

```

LOCATIONS: 1..789
US-09-649-866a-1 Ceres Seq. ID 1457851
Query Match 99.54; Score 788.6; DB 25; Length 789;
Best Local Similarity 100.0%; Pred. No. 2,3e-199;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
1 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
DB 1 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 61 TGAAGAGATGAGTTCTACACGAAAGATGACAGTGGCAGTGGAGATCAGACCGTA 120
DB 61 TGAAGAGATGAGTTCTACACGAAAGATGACAGTGGCAGTGGAGATCAGACCGTA 120
QY 121 GAGGCAATTAAGACCACTAGTCTTTGTGCGTGAATCTACTACTCCGTCGCTTAAT 150
DB 121 GAGGCAATTAAGACCACTAGTCTTTGTGCGTGAATCTACTACTCCGTCGCTTAAT 150
QY 181 CAACATCTCCGGAACAAGTTAGATCTGTTCTCAAGGAGAAAGTTCTCTCTCTCT 240
DB 181 CAACATCTCCGGAACAAGTTAGATCTGTTCTCAAGGAGAAAGTTCTCTCTCTCT 240
QY 241 GTCTCCGAGCCGTTACTCTCTCTGCTGAGAGCGAAGACGAAAGAACTTTTCCT 300
DB 241 GTCTCCGAGCCGTTACTCTCTCTGCTGAGAGCGAAGACGAAAGAACTTTTCCT 300
QY 301 GAGAAACATTTGATCAGAGCTTTTAAAGAAAGATGGAATTCACCGCAGAGCGTTAAG 360
DB 301 GAGAAACATTTGATCAGAGCTTTTAAAGAAAGATGGAATTCACCGCAGAGCGTTAAG 360
QY 361 CAATACCAACGACCGAGCAATCAAGCCGATCTAGTCCCAATCAAGCCGCTGTCTC 420
DB 361 CAATACCAACGACCGAGCAATCAAGCCGATCTAGTCCCAATCAAGCCGCTGTCTC 420
QY 421 GAYGTCCACACCGGAAATCCTTCTACAGGCTCCGAGGCGATTAATGATGTTCCCGGA 480
DB 421 GAYGTCCACACCGGAAATCCTTCTACAGGCTCCGAGGCGATTAATGATGTTCCCGGA 480
QY 481 AAAGACGACGACGAGCTTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
DB 481 AAAGACGACGACGAGCTTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
QY 541 CTGGAAGCTCTCAGTGAAGAAAGATCAATCTTAATGATTTGGAGACCAATTTGA 600
DB 541 CTGGAAGCTCTCAGTGAAGAAAGATCAATCTTAATGATTTGGAGACCAATTTGA 600
QY 601 GCTAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 GCTAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GTTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 ATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 TTTGCTGTT 789
DB 781 TTTGCTGTT 789

```


Qy 775 CCGGAGATTGCGT 789
 Db 601 CCGGAGATTGCGT 615

RESULT 6

US-09-513-996A-12108
 ; Sequence 12108, Application US/09513996A
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-709P
 ; CURRENT APPLICATION NUMBER: US/09/513,996A
 ; NUMBER OF SEQ. ID NOS: 81028
 ; SEQ. ID NO 12108
 ; LENGTH: 522
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: 1..522
 ; OTHER INFORMATION: any n or Xaa = unknown
 ; FEATURE:
 ; OTHER INFORMATION: Location 1..522 / Ceres Seq. ID 1376054
 US-09-513-996A-12108

Query Match

Best Local Similarity 62.0%; Score 489.2; DB 19; Length 522;
 Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 294 TTCCTTGAAGCAATGATCAGAGCTTTAAGAAAAGATGAGATTCACCGCAGAGCA 353
 Db 14 TTCTTCCAGAACATATGATCAGAGCTTTAAGAAAAGATGAGATTCACCGCAGAGCA 73
 Qy 354 GCTAACCCATACAAAGCAGACCGAATCAAGCCGATCTACGTCGCAATCAAAAGCCG 413
 Db 74 GCTAACCCATACAAAGCAGACCGAATCAAGCCGATCTACGTCGCAATCAAAAGCCG 133
 Qy 414 TGTGTCGATGTCACACCGGAAAAATCCTTCTACAGGCTCCGAGAGCATTCGATGTT 473
 Db 134 TGTGTCGATGTCACACCGGAAAAATCCTTCTACAGGCTCCGAGAGCATTCGATGTT 193
 Qy 474 CGCGGAAAAAGACGAGACAGCTTGGGTAGATGATGAAGAACGAGAAGATGTGTC 533
 Db 194 CGCGGAAAAAGACGAGACAGCTTGGGTAGATGATGAAGAACGAGAAGATGTGTC 253
 Qy 534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGATCAATCTTTAATGATTGGAGACCA 593
 Db 254 TCCTTCTCTTGAAGGTCTCACTGAGAAAGATCAATCTTTAATGATTGGAGACCA 313
 Qy 594 ATTTGAAGCTAAGTATCGTGCTGGCCGCTGTCTCTTAAGGCTCTCTTGAAGTT 653
 Db 314 ATTTGAAGCTAAGTATCGTGCTGGCCGCTGTCTCTTAAGGCTCTCTTGAAGTT 373
 Qy 654 GCACATATGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 713
 Db 374 GCACATATGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 433
 Qy 714 TGTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 773
 Db 434 TGTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 493
 Qy 774 TCGGAGATTGCGT 789
 Db 494 TCGGAGATTGCGT 509

RESULT 7

US-09-620-394B-328
 ; Sequence 328, Application US/09620394B

GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai
 APPLICANT: BEVBERG, Vyacheslav
 TITLE OF INVENTION: Sequence-determined DNA fragments and corresponding polypeptides
 FILE REFERENCE: 2750-1067P
 CURRENT APPLICATION NUMBER: US/09/620,394B
 NUMBER OF SEQ. ID NOS: 9131
 ; SEQ. ID NO 328
 ; LENGTH: 522
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..522
 ; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
 ; NAME/KEY: misc feature
 ; LOCATION: 1..522
 ; OTHER INFORMATION: Ceres Seq. ID 1376054
 US-09-620-394B-328

Query Match

Best Local Similarity 62.0%; Score 489.2; DB 24; Length 522;
 Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 294 TTCCTTGAAGCAATGATCAGAGCTTTAAGAAAAGATGAGATTCACCGCAGAGCA 353
 Db 14 TTCTTCCAGAACATATGATCAGAGCTTTAAGAAAAGATGAGATTCACCGCAGAGCA 73
 Qy 354 GCTAACCCATACAAAGCAGACCGAATCAAGCCGATCTACGTCGCAATCAAAAGCCG 413
 Db 74 GCTAACCCATACAAAGCAGACCGAATCAAGCCGATCTACGTCGCAATCAAAAGCCG 133
 Qy 414 TGTGTCGATGTCACACCGGAAAAATCCTTCTACAGGCTCCGAGAGCATTCGATGTT 473
 Db 134 TGTGTCGATGTCACACCGGAAAAATCCTTCTACAGGCTCCGAGAGCATTCGATGTT 193
 Qy 474 CGCGGAAAAAGACGAGACAGCTTGGGTAGATGATGAAGAACGAGAAGATGTGTC 533
 Db 194 CGCGGAAAAAGACGAGACAGCTTGGGTAGATGATGAAGAACGAGAAGATGTGTC 253
 Qy 534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGATCAATCTTTAATGATTGGAGACCA 593
 Db 254 TCCTTCTCTTGAAGGTCTCACTGAGAAAGATCAATCTTTAATGATTGGAGACCA 313
 Qy 594 ATTTGAAGCTAAGTATCGTGCTGGCCGCTGTCTCTTAAGGCTCTCTTGAAGTT 653
 Db 314 ATTTGAAGCTAAGTATCGTGCTGGCCGCTGTCTCTTAAGGCTCTCTTGAAGTT 373
 Qy 654 GCACATATGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 713
 Db 374 GCACATATGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 433
 Qy 714 TGTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 773
 Db 434 TGTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 493
 Qy 774 TCGGAGATTGCGT 789
 Db 494 TCGGAGATTGCGT 509

RESULT 8

US-09-534-859-278.C
 ; Sequence 278, Application US/09534859

GENERAL INFORMATION:
 APPLICANT: Bush, David F.
 APPLICANT: Last, Robert L.
 APPLICANT: Levin, Irena M.
 APPLICANT: Norris, Susan R.
 APPLICANT: Purcell, Laurence D.
 APPLICANT: Reimschuessel, Steven D.

```
APPLICANT: Miegand, Roger C.
TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
FILE REFERENCE: 38-10(15493)B
CURRENT APPLICATION NUMBER: US/09/534,859
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 1127
SEQ ID NO 278
LENGTH: 103495
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-534-859-278

Query Match
Best Local Similarity: 59.0%; Score 489.2; DB 20; Length 103495;
Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 294 TTCCTTGAGAAACAAATTGATGAGCTTTAAAGAAAAAGATGAAATTCACCGGAGACA 353
DB 57585 TTCTTCCAGAAACAAATTGATGAGCTTTAAAGAAAAAGATGAAATTCACCGGAGACA 57526
QY 354 GCTAACCCATATCAAGCGACCGACGATCAAGCCGATCTACGTCGCAATCAAGGCCG 413
DB 57525 GCTAACCCATATCAAGCGACCGACGATCAAGCCGATCTACGTCGCAATCAAGGCCG 57466
QY 414 TGTGTGCAATGTCACCCAGCGAAATCCTTCTACGCGCTCCGAGCGCATTTACTGATGTT 473
DB 57465 TGTGTGCAATGTCACCCAGCGAAATCCTTCTACGCGCTCCGAGCGCATTTACTGATGTT 57406
QY 474 CCGCGAAAAAGCGACGACGAGCTTTGGGTAAGATGAGTAAGAAAGAAAGATGTGTC 533
DB 57405 CCGCGAAAAAGCGACGACGAGCTTTGGGTAAGATGAGTAAGAAAGAAAGATGTGTC 57346
QY 534 TCCTTCTCTGAGAGTCTCACTGAGAAAGATCAATCTTAATGATGGGAGACCA 593
DB 57345 TCCTTCTCTGAGAGTCTCACTGAGAAAGATCAATCTTAATGATGGGAGACCA 57286
QY 594 ATTGAAGCTAATATCTGCTGCTGGCCGCTGTGTCTCTTGAAGTCTCTCTCTGAGAT 653
DB 57285 ATTGAAGCTAATATCTGCTGCTGGCCGCTGTGTCTCTTGAAGTCTCTCTCTGAGAT 57226
QY 654 GCACATATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 713
DB 57225 GCACATATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 57166
QY 714 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 773
DB 57165 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 57106
QY 774 TCGGGGATTTGCTGTT 789
DB 57105 TCGGGGATTTGCTGTT 57090

RESULT 9
US-09-803-736-278/c
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Levin, Irene M.
APPLICANT: Norris, Susan R.
APPLICANT: Rounsley, Steven D.
APPLICANT: Miegand, Roger C.
TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
FILE REFERENCE: 38-10(15493)D
CURRENT APPLICATION NUMBER: US/09/803,736
CURRENT FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 09/534,859
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: identified by Attorney Docket number 04983, 0206CFUS01 38-10
NUMBER OF SEQ ID NOS: 1582
SEQ ID NO 278
LENGTH: 103495
```

```
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-803-736-278

Query Match
Best Local Similarity: 59.0%; Score 489.2; DB 31; Length 103495;
Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 294 TTCCTTGAGAAACAAATTGATGAGCTTTAAAGAAAAAGATGAAATTCACCGGAGACA 353
DB 57585 TTCTTCCAGAAACAAATTGATGAGCTTTAAAGAAAAAGATGAAATTCACCGGAGACA 57526
QY 354 GCTAACCCATATCAAGCGACCGACGATCAAGCCGATCTACGTCGCAATCAAGGCCG 413
DB 57525 GCTAACCCATATCAAGCGACCGACGATCAAGCCGATCTACGTCGCAATCAAGGCCG 57406
QY 414 TGTGTGCAATGTCACCCAGCGAAATCCTTCTACGCGCTCCGAGCGCATTTACTGATGTT 473
DB 57465 TGTGTGCAATGTCACCCAGCGAAATCCTTCTACGCGCTCCGAGCGCATTTACTGATGTT 57406
QY 474 CCGCGAAAAAGCGACGACGAGCTTTGGGTAAGATGAGTAAGAAAGAAAGATGTGTC 533
DB 57405 CCGCGAAAAAGCGACGACGAGCTTTGGGTAAGATGAGTAAGAAAGAAAGATGTGTC 57346
QY 534 TCCTTCTCTGAGAGTCTCACTGAGAAAGATCAATCTTAATGATGGGAGACCA 593
DB 57345 TCCTTCTCTGAGAGTCTCACTGAGAAAGATCAATCTTAATGATGGGAGACCA 57286
QY 594 ATTGAAGCTAATATCTGCTGCTGGCCGCTGTGTCTCTTGAAGTCTCTCTCTGAGAT 653
DB 57285 ATTGAAGCTAATATCTGCTGCTGGCCGCTGTGTCTCTTGAAGTCTCTCTCTGAGAT 57226
QY 654 GCACATATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 713
DB 57225 GCACATATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 57166
QY 714 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 773
DB 57165 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 57106
QY 774 TCGGGGATTTGCTGTT 789
DB 57105 TCGGGGATTTGCTGTT 57090

RESULT 10
US-09-654-617-126256
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jinsong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/654,617
CURRENT FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 126256
LENGTH: 656
TYPE: DNA
ORGANISM: Arabidopsis thaliana
OTHER INFORMATION: unsure at all n locations
US-09-654-617-126256

Query Match
Best Local Similarity: 60.7%; Score 478.6; DB 25; Length 656;
Matches 478; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 311 TGTATGAGCTTTAAAGAAAGATGAGATTCACCGACGACGATTAAGCAATACGACG 370
DB 40 TGTATGAGCTTTAAAGAAAGATGAGATTCACCGACGACGATTAAGCAATACGACG 99
QY 371 GCACCGAGCATCAAGCGCATCTACGTCGCAATCAAGCGCGTGTGTGAGATGACCA 430
```

DB 100 GCACGACGAAATCAAAAGCCCATCTACGTCGCAATCAAAAGCCGCTGTCTTGCATGTCACCA 159
CY 431 CUGAAATCTCTTACGGCTCCGAGGCGATTAATCGATGTCGCGGAAAGACGGGA 490
DB 160 CCGGAAATCTCTTACGGCTCCGAGGCGATTAATCGATGTCGCGGAAAGACGGGA 219
CY 491 GCAGAGCTTTGGGTAGATAGTAAAGCAAGAGATGTCTCTCTCTCTTGAAGTTC 550
DB 220 GCAGAGCTTTGGGTAGATAGTAAAGCAAGAGATGTCTCTCTCTCTTGAAGTTC 279
CY 551 TCACGTGAAAGAGATCAATCTCTTAATGATGAGACCAAAATTTGAAGCTAGATTC 610
DB 280 TCACGTGAAAGAGATCAATCTCTTAATGATGAGACCAAAATTTGAAGCTAGATTC 339
CY 611 CTGTGCTGGCCGCTGTCTCTTACGATCTCTCTTGAAGTGAATGATGATTAAC 670
DB 340 CTGTGCTGGCCGCTGTCTCTTACGATCTCTCTTGAAGTGAATGATGATTAAC 399
CY 671 TATTGTGTGAGAGATCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 730
DB 400 TATTGTGTGAGAGATCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
CY 731 TTGTGATCAATTAACCAATAGTACCAATTAATTAATTAATTAATTAATTAATTAATTA 789
DB 460 TTGTGATCAATTAACCAATAGTACCAATTAATTAATTAATTAATTAATTAATTAATTA 518

RESULT 11

US-09-684-016-126256
Sequence 126256, Application US/09684016
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT FILING DATE: US/09/684, 016
CURRENT APPLICATION NUMBER: 2000-10-10
PRIORITY FILING DATE: 2000-09-05
PRIORITY APPLICATION NUMBER: 2000-09-05
SEQUENCE OF SEQ ID NOS: 463173
LENGTH: 656
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: unsure
LOCATION: (1) (656)
OTHER INFORMATION: unsure at all n locations
US-09-684-016-126256

Query Match 60.7%; Score 478.6; DB 27; Length 656;
Best Local Similarity 99.8%; Pled. No. 1,1e-116;
Matches 478; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 311 TGATGAGCTTTAAAGAAAGATGGAATTCACCGGAGAGAGCTTAAGCCATTAACG 370
DB 40 TGATGAGCTTTAAAGAAAGATGGAATTCACCGGAGAGAGCTTAAGCCATTAACG 59
CY 371 GCACGACGAATCAAAAGCCGATCTACGATCAAAAGCCGCTGTGTGTGTGTGTGTGTGT 430
DB 100 GCACGACGAATCAAAAGCCGATCTACGATCAAAAGCCGCTGTGTGTGTGTGTGTGTGT 159
CY 431 CCGGAAATCTCTTACGGCTCCGAGGCGATTAATGATGTCGCGGAAAGACGCGA 490
DB 160 CCGGAAATCTCTTACGGCTCCGAGGCGATTAATGATGTCGCGGAAAGACGCGA 219
CY 491 GCAGAGCTTTGGGTAGATAGTAAAGCAAGAGATGTCTCTCTCTCTTGAAGTTC 550
DB 220 GCAGAGCTTTGGGTAGATAGTAAAGCAAGAGATGTCTCTCTCTCTTGAAGTTC 279
CY 551 TCACGTGAAAGAGATCAATCTCTTAATGATGAGACCAAAATTTGAAGCTAGATTC 610
DB 280 TCACGTGAAAGAGATCAATCTCTTAATGATGAGACCAAAATTTGAAGCTAGATTC 339

CY 611 CTGTGCTGGCCGCTGTCTCTTACGATCTCTCTTGAAGTGAATGATGATTAAC 670
DB 340 CTGTGCTGGCCGCTGTCTCTTACGATCTCTCTTGAAGTGAATGATGATTAAC 399
CY 671 TATTGTGTGAGAGATCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 730
DB 400 TATTGTGTGAGAGATCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
CY 731 TTGTGATCAATTAACCAATAGTACCAATTAATTAATTAATTAATTAATTAATTAATTA 789
DB 460 TTGTGATCAATTAACCAATAGTACCAATTAATTAATTAATTAATTAATTAATTAATTA 518

RESULT 12

US-09-565-309A-8446
Sequence 8446, Application US/09565309A
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 3750-0853P
CURRENT FILING DATE: US/09/565, 309A
CURRENT APPLICATION NUMBER: 2000-05-05
SEQUENCE OF SEQ ID NOS: 68449
LENGTH: 521
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) (521)
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
LOCATION: (1) (521)
OTHER INFORMATION: 10261:4974 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-8446

Query Match 60.5%; Score 477.2; DB 22; Length 521;
Best Local Similarity 98.8%; Pled. No. 2,4e-116;
Matches 490; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

CY 294 TTCCCTTGAAACCAATGATGAGAGCTTTAAAGAAAGATGGAATTCACCGGAGAGCA 353
DB 14 TTCCCTTGAAACCAATGATGAGAGCTTTAAAGAAAGATGGAATTCACCGGAGAGCA 73
CY 354 GTTAAGCCATTAACCAAGCCGACGAGATCAAAAGCCGATCTACGATCAAAAGCCG 413
DB 74 GTTAAGCCATTAACCAAGCCGACGAGATCAAAAGCCGATCTACGATCAAAAGCCG 133
CY 414 TTTGTGAGATCTACCAAGCCGAGATCTCTTACGATCTCCGAGAGAGATGATG 473
DB 134 TTTGTGAGATCTACCAAGCCGAGATCTCTTACGATCTCCGAGAGAGATGATG 192
CY 474 CCGGAAATCTCTTACGGCTCCGAGGCGATTAATGATGTCGCGGAAAGACGCGA 533
DB 193 CCGGAAATCTCTTACGGCTCCGAGGCGATTAATGATGTCGCGGAAAGACGCGA 252
CY 534 TCCTTCTCTTAAAGGCTACGAGAAAGATCAATCTCTTAATGATGAGACCA 593
DB 253 TCCTTCTCTTAAAGGCTACGAGAAAGATCAATCTCTTAATGATGAGACCA 312
CY 594 ATTGAAGCTAGATTAATGATGAGAGCTTTGAGGCGATGATCTCTCTCTCTGAGAT 653
DB 313 ATTGAAGCTAGATTAATGATGAGAGCTTTGAGGCGATGATCTCTCTCTCTGAGAT 372
CY 654 GCACGTATGATGATTAATGATGAGAGCTTTGAGGCGATGATCTCTCTCTCTGAGAT 713
DB 373 GCACGTATGATGATTAATGATGAGAGCTTTGAGGCGATGATCTCTCTCTCTGAGAT 432
CY 714 TTTGTGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773

Db 433 TGTTCGATCTGATCGTTTGTATACATTACATAGTACCAATTAATCATGAATAA 492
QY 774 TCGGGATTCGTGTT 789
Db 493 TCGGGATTCGTGTT 508

RESULT 13
US-09-565-309A-43324

/ Sequence 43324, Application US/09565309A
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai
/ APPLICANT: BROVER, Vyacheslav
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-0853P
/ CURRENT APPLICATION NUMBER: US/09/565,309A
/ CURRENT FILING DATE: 2000-05-05
/ NUMBER OF SEQ ID NOS: 68449
/ SEQ ID NO 43324
/ LENGTH: 521
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(521)
/ OTHER INFORMATION: any n = a, g, c, t, unknown, or other
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(521)
/ OTHER INFORMATION: 10261 : 5TAG CONSENSUS (Clone Number:5tag_consensus)
US-09-565-309A-43324

Query Match

Best Local Similarity 60.5%; Score 477.2; DB 22; Length 521;
Best Local Similarity 98.8%; Pred. No. 2.4e-115;
Matches 490; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 294 TTCCTGTGAAACATTTATGATGAGCTTTAAAGAAAAGATGGAATTCACCGAGCA 353
Db 14 TTTTTCAGAAACATTTATGATGAGCTTTAAAGAAAAGATGGAATTCACCGAGCA 73
QY 354 GCTAACCAATACAAAGGACCGACCAATCAAGCCGATCTACGTCGCAATCAAGCCG 413
Db 74 GCTAACCAATACAAAGGACCGACCAATCAAGCCGATCTACGTCGCAATCAAGCCG 133
QY 414 TGTTCGATGTCACACCGGAAATTCCTTCTACGCTCCGAGGCCATTCGTGATGT 473
Db 134 TGTTCGATGTCACACCGGAAATTCCTTCTACGCTCCGAGGCCATTCGTGATGT 192
QY 474 CCGCGAAAGAGCGAGAGAGCTTTGGGTAGATGATGAAGAAAGAAAGATGTGTG 513
Db 193 CCGCGAAAGAGCGAGAGAGCTTTGGGTAGATGATGAAGAAAGAAAGATGTGTG 252
QY 534 TCTTCTCTTGAAGGTCTCTACCTGAGAAAGATCAATCTTATGATTTGGAGACCA 553
Db 253 TCTTCTCTTGAAGGTCTCTACCTGAGAAAGATCAATCTTATGATTTGGAGACCA 312
QY 594 ATTGAAGTACGATTCCTGCTGGCGGCTGTCTCTTAGGCTCTCTTGTGAGATT 653
Db 313 ATTGAAGTACGATTCCTGCTGGCGGCTGTCTCTTAGGCTCTCTTGTGAGATT 372
QY 654 GCACTATGATATGATCTATTTGTGTGAGAGATCTTGTGTGTGTGTGTGTGTG 713
Db 373 GCACTATGATATGATCTATTTGTGTGAGAGATCTTGTGTGTGTGTGTGTGTG 452
QY 714 TGTTCGATCTGATCGTTTGTATACATTACATAGTACCAATTAATCATGAATAA 773
Db 433 TGTTCGATCTGATCGTTTGTATACATTACATAGTACCAATTAATCATGAATAA 492
QY 774 TCGGGATTCGTGTT 789
Db 493 TCGGGATTCGTGTT 508

RESULT 14
US-09-770-961-478

/ Sequence 478, Application US/09770961
/ GENERAL INFORMATION:
/ APPLICANT: Goriach, Jörn
/ APPLICANT: An, Yong-Oliang
/ APPLICANT: Hamilton, Carol M.
/ APPLICANT: Price, Jennifer L.
/ APPLICANT: Raines, Tracy M.
/ APPLICANT: Yu, Yang
/ APPLICANT: Ramekha, Joshua G.
/ APPLICANT: Page, Amy
/ APPLICANT: Matthew, Abraham V.
/ APPLICANT: Ledford, Brooke L.
/ APPLICANT: Kressner, Jeffrey P.
/ APPLICANT: Haas, William David
/ APPLICANT: Kitcher, Maya
/ APPLICANT: Slader, Ted
/ APPLICANT: Davis, Keith R.
/ APPLICANT: Allen, Keith
/ APPLICANT: Hoffman, Neil
/ APPLICANT: Harban, Patrick
/ TITLE OF INVENTION: Expressed Sequences of Arabidopsis
/ FILE REFERENCE: 2026 (PARA-015PRV)
/ CURRENT APPLICATION NUMBER: US/09/770,961
/ PRIOR APPLICATION NUMBER: 2001-01-26
/ PRIOR FILING DATE: 2000-01-27
/ NUMBER OF SEQ ID NOS: 999
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 478
/ LENGTH: 458
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-09-770-961-478

Query Match

Best Local Similarity 59.6%; Score 470.6; DB 30; Length 498;
Best Local Similarity 99.8%; Pred. No. 1.3e-114;
Matches 470; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 319 GCTTAAAGAAAGATGATTCACCGAGAGCACTAAGCAATACAGCGACCGAC 378
Db 10 GCTTAAAGAAAGATGATTCACCGAGAGCACTAAGCAATACAGCGACCGAC 69
QY 379 GATCAAGCCGATCTACCTGCAATCAAGGCGGTGTGTGATGATCAAGCGGAAA 438
Db 79 GATCAAGCCGATCTACCTGCAATCAAGGCGGTGTGTGATGATCAAGCGGAAA 129
QY 439 TCTTCTAGGCTCCGAGAGCACTTCTGATGTTGCGGAAAGAGCGAGAGACT 498
Db 130 TCTTCTAGGCTCCGAGAGCACTTCTGATGTTGCGGAAAGAGCGAGAGACT 189
QY 499 TTGGTAAAGTAAAGAAAGAAAGATGTGTCTTCTTGAAGGTCTCACTAG 558
Db 190 TTGGTAAAGTAAAGAAAGAAAGATGTGTCTTCTTGAAGGTCTCACTAG 249
QY 559 AAAGATCAATCTTATATGATTTGGAGACCAATTTGAAGCTATGATCTGTG 618
Db 250 AAAGATCAATCTTATATGATTTGGAGACCAATTTGAAGCTATGATCTGTG 309
QY 619 GCGCGTGTCTCTTATGATCTCTTCTGAGATGCACTATGATGATGATGATG 678
Db 310 GCGCGTGTCTCTTATGATCTCTTCTGAGATGCACTATGATGATGATGATG 369
QY 679 GTGAGATCTTG 738
Db 370 GTGAGATCTTG 429
QY 739 AATTACATAGTACCAATTAATCATGAATTAATCGGGSATTCGTGTT 789
Db 430 AATTACATAGTACCAATTAATCATGAATTAATCGGGSATTCGTGTT 480

```
RESULT 15
US-09-565-309A-1449
; Sequence 1449, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 1449
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(456)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: (1)..(456)
; OTHER INFORMATION: 1916:36727 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-1449

Query Match          57.7%; Score 455.6; DB 22; Length 456;
Best Local Similarity 99.8%; Pred. No. 1.3e-110;
Matches 455; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCATCAACAAAAAATTCATACACAAAAACAAAAACAAAAAGTTAATTC 60
   |||
Db 1 ATCATCAACAAAAAATTCATACACAAAAACAAAAACAAAAAGTTAATTC 60

QY 61 TGAAGAAAGATGATTTTACACAGCAAGCATGACAGTGGCATCGGAGCCGTA 120
   |||
Db 61 TGAAGAAAGATGATTTTACACAGCAAGCATGACAGTGGCATCGGAGCCGTA 120

QY 121 GAGGCATTTAAAGACCACTAGGCTTTGTCGGTGAACATCACTACCTCCGTCGTTAAT 180
   |||
Db 121 GAGGCATTTAAAGACCACTAGGCTTTGTCGGTGAACATCACTACCTCCGTCGTTAAT 180

QY 181 CAACATCTCCGGAACAAGTTAGATCTGTTCTCAAGGAAAAAGTTCTTCTCTCT 240
   |||
Db 181 CAACATCTCCGGAACAAGTTAGATCTGTTCTCAAGGAAAAAGTTCTTCTCTCTCT 240

QY 241 GTCTCCGAGCGCTTACCTCTCTGAGAGGAGAGCAAGCAAGCACTTTCCCTT 300
   |||
Db 241 GTCTCCGAGCGCTTACCTCTCTGAGAGGAGAGCAAGCAAGCACTTTCCCTT 300

QY 301 GAGAAACAAATGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCACTAAGC 360
   |||
Db 301 GAGAAACAAATGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCACTAAGC 360

QY 361 CAATCAACGCGACCGAAGCAATCAAGCCGATCTACCTCGCAATCAAGGCGGTGTTC 420
   |||
Db 361 CAATCAACGCGACCGAAGCAATCAAGCCGATCTACCTCGCAATCAAGGCGGTGTTC 420

QY 421 GAGTCAACGCGAAGCAATCTTCTACGCGCTCCGGA 456
   |||
Db 421 GAGTCAACGCGAAGCAATCTTCTACGCGCTCCGGA 456
```

Search completed: January 8, 2003, 15:19:42
Job time : 3650 secs

451 GAGACTCTGACTGACTGGAGTCTCAGTTCACTTCAAGTATCATCACGTGGC 504

RESULT 9

Query Match	6.9%;	Score 54.4;	DB 10;	Length 1890;
Best Local Similarity	54.1%;	Pred. No. 9.4e-06;		
Matches 159;	Conservative 1;	Mismatches 122;	Indels 12;	Gaps 2

Qy	337	GATTCACCGCGAAGAGAGGTATAGCCAAATACACGGCACCGAGATATCAAAAGCCATCTAC	398
Db	292	GACTTACACCCCCCGAGCTGGCGGCGCTTCGACGGGCTCCAGGAGCCC--GGCGATTACTC	348
Qy	397	GTCGCATCATAAAGCCGCGTGTGTTCAGTGTACACCGCGAAATCTCTTACGCGCTCCGGA	458
Db	349	ATGCCATCATCAACGGCAAGGTGTTCATGTATGACCAAAAGCCGCGAAATTTACGGCGCCGAG	408
Qy	457	GCGCATTTACTGCATGTTTCCGCCGGAAGAAACCGGACGAGAGCTTTGGGTAAATAGTAAAG	518
Db	409	GCGCGCATGGGGTCTTTTCTGTGAAAGATGTCATCCAGGGGACCTTGCCACATTTGGCTG	468
Qy	517	AACGAGAA-----GATGTGTCTCTTCTTTAAAGTCTCACTAGAGAAAGATC	567
Db	469	GATAGGAAACACTAAGGATGAGTACATACCTTTCTGACTCACTGCTGCCAGGAG	528
Qy	568	AATATCTTTATGATTGGGAGACCAATTTGAGAGCTAAGTATCTCTGTTGGC	621
Db	529	GAGACTCTGTGATCGAGGAGTCTAGTTCATCTTCAAGTATATCAACCTGGGC	582

RESULT 10

US-09-783-590-11410
Sequence 11410, Application US/09783590
Parent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16-2CI
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420, 856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346, 731
PRIOR FILING DATE: 1994-11-21

```

; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11410
; LENGTH: 415

```

Query Match	5.7%	Score 44.8;	DB 10;	Length 415;
Best Local Similarity	57.8%	Pred. No. 0.0025;		
Matches	93;	Conservative	1;	Mismatches 66;
			Indels 1;	Gaps 1

QY 336 GGAATTACACCCAGAGCGAGCGTAACCCATATCAACGGCACCAGCAATATAAAGCCGATCTA 335
 Dd 37 GAACTTACCCCTCCGCCGAGGTGGCGGCTTACGAGGCGTTCAGGGACCCGGGAGCATAACT 155
 QY 396 CGTCCGAATCCAAAGCCCTGTGTTCCGAYGCAACACCGGAAAAATCCCTTCTACGCTCCGG 455
 Dd 156 CATGGCCATCTACGGCAAGAGTGTTCGATGTGACCAAAAGGCCCCAAATTTCTCGGAGCCGA 215
 QY 456 AGCGGATTACTCGATGTTCGCCGGAAAAAGACGCGACAGAG 496
 Dd 216 GGGGCGTNAATGGGCTTTTGTGTGGAAAGAGGTGATCCAGG 256

RESULT 11

US-09-960-352-11743
Sequence 11743, Application US/0960352
Parent No. US20060137139n1
GENERAL INFORMATION:
APPLICANT: Warner, Wesley C.
APPLICANT: Tao, Wengsheng
APPLICANT: Ryatt, John C.


```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 78
; LENGTH: 2776
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-245-78

```

```

Query Match 7.9%; Score 62.2; DB 9; Length 2776;
Best Local Similarity 54.6%; Pred. No. 6,5e-08;
Matches 171; Conservative 1; Mismatches 129; Indels 12; Gaps 2;

```

```

Cy 320 CTTTAAAGAAAGATGATTCACCGGACGACCTAAGCAATACACGCGCGCAGC 379
Db 314 CTGCGATGAAGAGCGGACCTTCACCTTGGACACGCTGCGACATGACGCGCTTCGCA 373
Cy 380 AATCAAGCGGATCTACGTCGCAATCAAGCGCGTGTGTGATGATCACCACCGGAAAT 439
Db 374 ACC---GGCATCTGCTCGCGGATGGAAGTCTTGACGTACCAAGATCGCA 430
Cy 440 CCTTACCGCTCCGAGCGCATTAAGTGTGCGGAAAGACCGGACCGAGCTT 499
Db 431 AGTTCTACCGCGCGCGGATTCATATGGAATATTTGCTGTAGGAGATGCTCCAGAGGAC 490
Cy 500 TGGGTAGATGATGAAGACGAGA-----AGATGTCTCTCTCTCTTAAAGTTC 550
Db 491 TGGCCACATTTTGCCTAGTAAAGATGACCTTAGAGATGAATATGATGATCTCTCAAT 550
Cy 551 TCATGAGAAAGATCACTTAAATGATTTGGAGACCAATTTGAAGCTAAGTATC 610
Db 551 TGATGCAATACAAATGAGAGTGTTCAGAGATGGGAATGCATTTAAAGAAATATG 610
Cy 611 CTGTGCTTGGCGC 623
Db 611 ATTATGTAGCGAG 623

```

RESULT 7

```

US-09-923-876-197
; Sequence 197, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (lco)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT FILING DATE: US/09/923, 876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 197

```

```

; LENGTH: 253
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700156530H1
; LOCATION: 135
; OTHER INFORMATION: a, c, g, or other
US-09-923-876-197

```

```

Query Match 6.9%; Score 54.8; DB 10; Length 253;
Best Local Similarity 68.5%; Pred. No. 2.6e-06;
Matches 74; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

```

```

Cy 403 ATCAAGCGCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
Db 1 ATCAAGCGCGATCTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Cy 463 TACTGATGTTGCGCGGAAAGACGAGCAGAGCTTGGTAAATG 510
Db 61 TACGCGCTGTTCGCGGCAAAAGATGCCAGAGAGCTCTAGCGAATG 108

```

RESULT 8

```

US-10-164-871-1
; Sequence 1, Application US/10164871
; Patent No. US2002017194A1
; GENERAL INFORMATION:
; APPLICANT: Hlata, Yuichi
; TITLE OF INVENTION: STEROID HORMONE BINDING PROTEIN
; FILE REFERENCE: 06501-059001
; CURRENT FILING DATE: 2002-06-07
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US/09/565,808
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: WO/99/05010
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: JP/9/322376
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11...(585)
US-10-164-871-1

```

```

Query Match 6.9%; Score 54.4; DB 9; Length 588;
Best Local Similarity 54.1%; Pred. No. 5.1e-06;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;

```

```

Cy 317 GATTCACCGCAGACGCTAAGCAATACAGCAGCAATCAAGCGCATCTAC 396
Db 214 GACTTCACCCCGCGAGCTGCGCGCTTCAAGCGCGTCCAGAGCC---GGCATACTC 270
Cy 337 GTGCAATCAAGCGCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 456
Db 211 ATGCGCATACAGCGCAAGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 330
Cy 457 GCGGATTAAGTGTGTGCGGAAAGACGAGCAGAGCTTGGTAAATGATGATGATGATGAT 516
Db 331 GCGCGGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
Cy 517 AACGAGAA-----GATGTCTCTCTCTTGAAGGTTCTACGAGAAAGATC 567
Db 391 GATAGGAGCACTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
Cy 568 AATACTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621

```


and XhoI was ligated to modified Lambda FLX-1 vector (Garnier et al., submitted for publication) digested with BamHI and SalI. The clone is in a modified pBluescript vector. Please visit our web site (http://www.jsc.fiken.go.jp/e/Plant/index_english.htm) for further details.

FEATURES

Location/Qualifiers

1..439
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAF11-09-L22"
/clone_id="RAF11"
/dev_stage="Plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site 1: BamHI; Site 2: SalI; subjected to various treatments (dehydration, cold, high salt, ABA, heat and UV). Dark-grown plants"
ORIGIN

BASE COUNT 134 a 106 c 84 g 115 t

Query Match 55.1%; Score 434.4; DB 10; Length 439;
Best Local Similarity 99.3%; Pred. No. 1.3e-85;
Matches 435; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

351 GCAAGTAAAGCAATCAAGCGACCGAGATCAATCAAGCGATCTACGTCGCAATCAAGC 410
439 GCAAGTAAAGCAATCAAGCGACCGAGATCAATCAAGCGATCTACGTCGCAATCAAGC 380
411 CCGTGTTCAGTCAACGACCGGAAATCTTCTACGTCGCGAGGCGATCTACGAT 470
379 CCGTGTTCAGTCAACGACCGGAAATCTTCTACGTCGCGAGGCGATCTACGAT 320
471 GTTCCCGGAAAGCAAGCGACCGAGATCTTCTACGTCGAGTCAAGCAATCAAGATGT 530
319 GTTCCCGGAAAGCAAGCGACCGAGATCTTCTACGTCGAGTCAAGCAATCAAGATGT 260
531 GTTCTCTCTCTGAAAGGTCTCTACTGAGAAAGACATCAATCTTAATGATTGGAGAC 590
259 GTTCTCTCTCTGAAAGGTCTCTACTGAGAAAGACATCAATCTTAATGATTGGAGAC 200
591 CAATTTAGCTAGTATCTCTGCTGCTGGCCGCTGCTGCTCTTACGTCCTTTTGAG 650
199 CAATTTAGCTAGTATCTCTGCTGCTGGCCGCTGCTGCTCTTACGTCCTTTTGAG 140
551 ATTGACATGTATGTACTATGTGTGTGAGATCTTGTGTGTGTGTGTGTGTGTGTGT 710
139 ATTGACATGTATGTACTATGTGTGTGAGATCTTGTGTGTGTGTGTGTGTGTGTGT 80
471 TCGTGTTCAGTCAACGACCGGAAATCTTCTACGTCGCGAGGCGATCTACGAT 770
79 TCGTGTTCAGTCAACGACCGGAAATCTTCTACGTCGCGAGGCGATCTACGAT 20
771 AATTCGGGATTTCTGTGT 788
19 AATTCGGGATTTCTGTGT 2

RESULT 2
LOCUS A1996124/c 436 bp mRNA linear EST 08-SEP-1993
DEFINITION 701550133 A. thaliana, Columbia Col-0, inflorescence-2 Arabidopsis
ACCESSION A1996124
VERSION A1996124.1 GI:5843029
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicot;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopses.

REFERENCE
AUTHORS Chen J., Montoya M., Chan E., McCreary M., Garrison E., Gilibert J.,
Kang X., Hillman J., Guejter K., Kim C., Doyle M., Szostak P.,

Georgene G., Burns D., Griffin J., Mouanoutoua M., Nguyen D., Tan R.,
Rose M., Warren B., Ton B., Kastury K., Portillo C., Carpio T.,
Folick J., Suzuki G., Argentine C., Shah S., Nobrega A., Murry L.,
Turner C., Kirkorian S., Elder L. and Hanson D.
Arabidopsis thaliana Gene Expression MicroArray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63124, USA
Tel: 877-577-2733
Fax: 314-427-3734
E-mail: service@genomesystems.com.

FEATURES

Location/Qualifiers

1..435

/organism="Arabidopsis thaliana"
/cultivar="Columbia Col-0"
/db_xref="taxon:3702"
/clone="701550133"
/clone_id="A. thaliana, Columbia Col-0, inflorescence-2"
/tissue_type="inflorescence"
/dev_stage="4 - 7 weeks"
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated inflorescence tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

BASE COUNT 128 a 103 c 85 g 118 t
ORIGIN

Query Match 51.8%; Score 409; DB 9; Length 435;
Best Local Similarity 97.7%; Pred. No. 5.2e-80;
Matches 423; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

327 GAAAGAGATGAATTCACCGGACGACGCTAAGCAATAC-AACGGACCGAGCAATCAA 385
436 GAAAGAGATGAATTCACCGGACGACGCTAAGCAATACGCTAAGCAATCAA 377
386 AGCGATCTAGCTGCAATCAAGCGCGTGTGTGAGTCAACGACCGGAAATCTCTCT 445
376 AGCGATCTAGCTGCAATCAAGCGCGTGTGTGAGTCAACGACCGGAAATCTCTCT 317
446 AGCGATCTAGCTGCAATCAAGCGCGTGTGTGAGTCAACGACCGGAAATCTCTCT 505
316 AGCGATCTAGCTGCAATCAAGCGCGTGTGTGAGTCAACGACCGGAAATCTCTCT 257
506 AGTATGATGAAGCAAGATGTGTCTCTCTCTTGAAGTCTCACTGAGAAAGAGA 565
256 AGTATGATGAAGCAAGATGTGTCTCTCTCTTGAAGTCTCACTGAGAAAGAGA 197
566 TCAATCTCTTATGATGGAGACCAATTTAGTAACTAGTATCTCTGTTGGCGG 625
196 TCAATCTCTTATGATGGAGACCAATTTAGTAACTAGTATCTCTGTTGGCGG 137
626 TTGTCTCTTAAAGTCTCTCTTGAAGTCACTATGTTAAGTCACTATGTGTGAGGA 685
136 TTGTCTCTTAAAGTCTCTCTTGAAGTCACTATGTTAAGTCACTATGTGTGAGGA 77
685 TCTTGT 745
76 TCTTGT 17

RESULT 3

```

QY 320 CTTTAAAGAAAAAGATGATATTCACCCGAGACGCTAACCCATACACGGGACCGGACG 375
      |||
DB 287 CTCGCATGAAGAGGGGAGCTTTCAGCTTGAGAGCGAGCTGCGCCAGTACGAGCGCTCCCGCA 346
      |||
QY 380 AATCAAGCCGATCTACGTCGCATCAAAAGCCGCTGTTCGAVGTCCACCCGCAAAAT 439
      |||
DB 347 ACCC---GCGCATCTCTGCTCCGCGTCAATGGGAAGCTTCGACGTGACCAAGGAGCA 403
      |||
QY 440 CCTTACGCGCTCCGAGCGCATTACTCGATGTTCCCGGAAAAAGACCGAGAGAGCTT 499
      |||
DB 404 AGTCTACGCGCCGCGCGGCTCATATGGAATATTTGCTGTAGGGAATGCTCCAGAGAGCA 463
      |||
QY 500 TGGGTAAAGATGATAGAACGAGA-----AGATGTCTCTCTCTTCTGAAAGCTC 550
      |||
DB 464 TGGCCACATTTTTCCTAATAAAGATGACCTTAGAGATGATGATCTCTCAGATC 523
      |||
QY 551 TCACGTAGAAAGATCAATACCTTTAATGATGGGAGACCAATTGAAGCTAAGTATC 610
      |||
DB 524 TGAATGCAATGACAAATGAGAGGTGTGAGATGCGAAATGCAAGTTTAAAGAAAAATG 593
      |||
QY 611 CTGTGCTGGCCG 623
      |||
DB 584 ATTATGTAGGCGAG 596
      |||

```

RESULT 5

```

US-10-098-841-217
; Sequence 217, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qiang A.
; APPLICANT: Ren, Peiyuan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1 Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098, 841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598, 042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: PL_genes Version 1.0
; SEQ ID NO 217
; LENGTH: 1936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(752)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1936)
; OTHER INFORMATION: n = a,c,t,c or g
US-10-098-841-217
Query Match

```

7.9% Score 62.2; DB 9; Length 1936;

```

Best local similarity 54.6%; Pred. 33.54e-08;
Matches 171; Conservative 1; Mismatches 129; Indels 12; Gaps 2.
QY 320 CTTTAAAGAAAAAGATGATATTCACCCGAGACGCTAACCCATACACGGGACCGGACG 379
      |||
DB 287 CTCGCATGAAGAGGGGAGCTTTCAGCTTGAGAGCGAGCTGCGCCAGTACGAGCGCTCCCGCA 426
      |||
QY 380 AATCAAGCCGATCTACGTCGCATCAAAAGCCGCTGTTCGAVGTCCACCCGCAAAAT 439
      |||
DB 347 ACCC---GCGCATCTCTGCTCCGCGTCAATGGGAAGCTTCGACGTGACCAAGGAGCA 403
      |||
QY 440 CCTTACGCGCTCCGAGCGCATTACTCGATGTTCCCGGAAAAAGACCGAGAGAGCTT 499
      |||
DB 404 AGTCTACGCGCCGCGCGGCTCATATGGAATATTTGCTGTAGGGAATGCTCCAGAGAGCA 463
      |||
QY 500 TGGGTAAAGATGATAGAACGAGA-----AGATGTCTCTCTCTTCTGAAAGCTC 550
      |||
DB 464 TGGCCACATTTTTCCTAATAAAGATGACCTTAGAGATGATGATGATCTCTCAGATC 603
      |||
QY 551 TCACGTAGAAAGATCAATACCTTTAATGATGGGAGACCAATTGAAGCTAAGTATC 610
      |||
DB 524 TGAATGCAATGACAAATGAGAGGTGTGAGATGCGAAATGCAAGTTTAAAGAAAAATG 663
      |||
QY 611 CTGTGCTGGCCG 623
      |||
DB 664 ATTATGTAGGCGAG 676
      |||

```

RESULT 6

```

US-09-984-245-78
; Sequence 78, Application US/09984245
; Patent No. US20020165317A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131

```

CY	629	TCTCTTGAGTCTCTCTTCGACATGGATTGCATATGAATCACTATTTGGTATCGAATCATC	638
DB	273	TCTCTTGAGTCTCTCTTCGACATGGATTGCATATGAATCACTATTTGGTATCGAATCATC	332
CY	689	TTGTGTTGTGTGTTTTCTGTGATTTCTGTTTGGATCGTGTGATTCATCATTAACCATA	748
DB	333	TTGTGTTGTGTGTTTTCTGTGATTTCTGTTTGGATCGTGTGATTCATCATTAACCATA	392
CY	749	AGTCCAAATATCTATGAATTAATCCGGGATTTTCGTTT	789
DB	393	AGTCCAAATATCTATGAATTAATCCGGGATTTTCGTTT	433
RESULT 11			
BI417544		429 bp	mRNA linear EST JF-NCU-200
LOCUS	BI417544		
DEFINITION	JN587302r Lotus japonicus nodule library 5 and 7 week-old Lotus		
ACCESSION	BI417544		
VERSION	BI417544.1	GI:15188567	
KEYWORDS	EST.		
SOURCE	Lectus japonicus.		
ORGANISM	Lectus japonicus.		
REFERENCE			
AUTHORS	Colebatch,G., Freund,S., Trevasakis,B and Urdavari,M.		
TITLE	Lectus japonicus root nodule ESTs: tools for functional genomics		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Urdavari MK Molecular Plant Nutrition Max Planck Institute of Molecular plant Physiology Am Muehlenberg 1, 14476 Golm, Germany Fax: 49 331 567 8250 Email: urdavarid@mpimp-golm.mpg.de Seq primer: T7 High quality sequence stop: 429. Location/Qualifiers 1. .429 /organism="Lotus japonicus" /cultivar="glfr (B-129)" /db_xref="taxon:34305" /clone_lib="Lotus japonicus nodule library 5 and 7 week-old" /dev_stage="5 and 7 week-old plants" /note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain K/A."		
FEATURES			
source			
EASE COUNT	114 a	118 c	109 g 88 t
ORIGIN			
Query Match	26.6%	Score 209.8;	DB 13; Length 429;
Best Local Similarity	79.4%;	Pred. No. 3.9e-36;	
Matches 247;	Conservative 1;	Mismatches 63;	Indels 0; Gaps 0
CY	325	AAAAAAGATGGAATTACCCGAGAGCACTAACGCAATACAACGCCGACGACGATCA	384
DB	79	AAGCAACAAGATGAGATGACCCGACAGCAACTAGACCAATACAACGCCGACGACCATCG	138
CY	385	AAGCGGATCTAACGTCGCAATCAAGAAGCCGCTGTCTTCGAYTCACCAACCGGAAAATCTTTC	444
DB	139	AAGCCAATCTACGCTCTCGTAGAGGCCCGCGTTCGATGATCACACCGGAAAATCTTTC	198
CY	445	TACGAGCTCCGAGGCGATTACTGATGATTTCCGCCGGAAGAAAGACGCGAGCAGGCTTGGGT	504
DB	159	TACGAGCCCCGGGTGGCGCTTACCGGATGTTTCCGCCGGAAGAGACCGCACAGAGCGCTTACCG	258
CY	505	AAGATAGCTAAGAACGAAAGATGTGTCTCTCTCTTTGAAGGTGTCTCTGAGAAAAG	564

D6	299	AATGTAAGTAAATGAACGAGGAGGAGGATTTCTTCGCCCGCAACTGGCATATGAGTTCTCCCAACAATTATG	338
OY	565	ATAAATACGTCTTATGATTTGGGAGACCAGAAATTGAAGCCTAAGAATCTCTTGCTTGCACCGCT	634
D6	313	ATCCGAGTCTTCTATCTACTGAGGAGAACAAATGCTTACTACTAAGTAACTCCTGTGTGCTCCG	378
OY	625	GTTGCTCTCTTA	635
D6	379	CTTCTTTCTTA	389

RESULT 13	
LOCUS	BQ791458
DEFINITION	E4375 Chinese cabbage etiolated seedling library Brassica rapa
ACCESSION	BQ791458
VERSION	BQ791458.1 GI:22006420
KEYWORDS	EST.
SOURCE	Brassica rapa subsp. pekinensis.
ORGANISM	Brassica rapa subsp. pekinensis Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica. Ryu,S.H., Yang,K.A., Lee,S.Y., Kim,H.-I., Cho,M.J. and Lim,C.O. Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNAs (2002)
REFERENCE	Ryu,S.H., Yang,K.A., Lee,S.Y., Kim,H.-I., Cho,M.J. and Lim,C.O. (2002)
AUTHORS	Unpublished (2002) Contact: Lim, C.O. Plant Molecular Biology & Biotechnology Research Centre Gyeongsang National University #500 Gajwa-dong, Jinju 660-701, Korea Tel: 82 55 751 6255 Fax: 82 55 759 9163 Email: colim@nongae.gsnu.ac.kr Seq primer: 17, location/Qualifiers source 1..344 /organism="Brassica rapa subsp. pekinensis" /cultivar="Jangwon" /db_xref="taxon:31351" /clone="E4375" /clone_1fb="Chinese cabbage etiolated seedling library" /tissue_type="Etiolated seedling" /lab_host="XL-1 Blue" /note="vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"
FEATURES	
CDS	1..344
BASE COUNT	96 a 97 c 86 g 65 t
ORIGIN	

Query Match	
Best Local Similarity 86.8%; Pred.No.1,je-33;	
Matches 217; Conservative 1; Mismatches 32; Indels 0; Gaps 0.	

OY	329	TAAATGAAATTCACCGGAGAGGAGCTATTAACCATTCACACGGCACCGACGAATCAAAGC	338
D6	95	ACGAATAGGATTCACCGGAGAGGAGCTATTAACCATTCACACGGACCGACTCATCAAGC	134
OY	389	CGATCTACGTGCAATCAAAGGCCGTGTGTGTCAGAYGCACACCGGAAATCTCTTAGC	443
D6	155	CGATCTACGTGCAATCAAAGGCCGTGTGTGTCAGAYGCACACCGGAAATCTCTTAGC	214
OY	443	GCTCCGAGGCGATTACTGATGTTGCCCGAAAAAGACCGAGAGAGGCTTGCGTAAAG	503
D6	215	GCGTAGAGCGCATTAACGAGATGTTCCGGGAAAAGACCGAGAGAGGCTTGCGGAAAG	274
OY	509	TGAGTAAGAAACGAAGATGTGTCTCTTCTTTGAAGGCTTCACGAGAAAGAGATCA	568
D6	275	TGAGCAAGAAACGAAGAGCGTGTCTCTTCTTCGAAATCTTCACGAGAAAGAGATCA	334
OY	569	ATACTCTTA	578
D6	335	CCACTCTTA	344

adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Ulla Vodka lab University of Illinois)."

E. T COUNT 134 a 153 c 110 g 115 t

Library Match

25.1%; Score 198.4; DB 14; Length 512;

Local Similarity 68.2%; Pred. No. 1.2e-33;

Matches 288; Conservative 1; Mismatches 132; Indels 1; Gaps 1;

```
C 328 AAAAAGATGCAATTCACCGCAGACGACTTAAGCCATATACACCGCAGCCGACGATCAAG 337
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
T 61 AAGACAAATGGAGCTGACACCCCGACAGCTGAGCCAAATACACCGCACCGACCATGAG 122
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 388 CCGATCTACGTGCAATCAAGGCCGCTGCTGAGTGCACCAACCGAATAATCCTTTAC 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
E 121 CCGATCTACGTGCGGTGAAGGCCGCTGCTGAGTGCACCAACCGAATAATCCTTTAC 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 448 GGCTCCGAGGCGCATTAAGTATGCTGCGGAAAGACCGACAGACCTTTGGTAAAG 507
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
T 181 GGCCCCGCGCGCCCTTACGCGCATGTTGCGCGCAAGACGCCACGACGAGCCCTGGCGAAG 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 508 ATGAGTAGAAGCAAGAAAGATGCTCTCTCTCTGAGGTCTCACTGAGAAAGATC 567
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
T 241 ATGAGCAAGAAAGCAAGACATCTCCCTCCCTGACGCGCTCTCCGACAAAGAGATC 302
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 568 AATACCTTAAATGAGAGACCAATTTGAGCTAAGATCTGCTGCTTGGCCGTGT 627
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
E 301 GCGCTCTCAACGACTGGAGAAACAAATTCACAGCTAAGACCTGTGCTGTCT 362
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 628 GTCTTTAGTCTCT-CTTGAAGATTGCACTATGTAATTAATTTGTGTGAGAT 686
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
T 361 CTCATTTAATCATATGATTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACT 422
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 687 CTTTGTTGTGTTTCTTTGATTTGTTGATTTGTTGATTTGTTGATTTGTTGATTTACCA 746
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
E 421 CTCCTTCTGTAACCGAAGTCAATCAACCTTTAATCCTTAATTAATTAATTAATTTCTTC 482
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
C 747 TA 748
      |||
T 481 TA 482
```

Search completed: January 8, 2003, 17:44:39
Time: 2154 secs

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130691.
PR 28-APR-1999; 99US-0131649.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143477.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 15-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158322.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

```

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 99.9%; Score 788.6; Da 21; Length 789;
Best Local Similarity 100.0%; Pred. No. 7, 1e-226;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

CY 1 ATCATCAACAAACAAATTCATATACAAACAAACAAACAAAGATTATTC 60
CB 1 ATCATCAACAAACAAATTCATATACAAACAAACAAACAAAGATTATTC 60
CY 121 GAGGATTTAAAGACCAACTAGTCTTTGTGAGTGAACATCACTCCGGTGAAT 180
CB 61 TGAAGAAAGATGAGTTCTACAGCAAGACATGACAGTGCAGTGCATCGAGCG 120
CY 121 GAGGATTTAAAGACCAACTAGTCTTTGTGAGTGAACATCACTCCGGTGAAT 180
CB 121 GAGGATTTAAAGACCAACTAGTCTTTGTGAGTGAACATCACTCCGGTGAAT 180
CY 181 CAACATCTCCGGAACAAAGATTAGATCTTTCTAAGGAAAGAGTTCTTCTTCT 240
CB 181 CAACATCTCCGGAACAAAGATTAGATCTTTCTAAGGAAAGAGTTCTTCTTCT 240
CY 241 GTCTCCGAGCGGTTACTCTCTGTGTGAGAGCAGAAAGAGAAAGCTTTCCCT 300
CB 241 GTCTCCGAGCGGTTACTCTCTGTGTGAGAGCAGAAAGAGAAAGCTTTCCCT 300
CY 301 GAGAAACAAATGATGACGCTTTAAAGAAAGATGGAATTCACCCGAGACTAAG 360
CB 301 GAGAAACAAATGATGACGCTTTAAAGAAAGATGGAATTCACCCGAGACTAAG 360
CY 361 CAATCAACGCGACGACGATCAAGCCGATCTAGCTCCCATTCAAAGCGCTGTTC 420
CB 361 CAATCAACGCGACGACGATCAAGCCGATCTAGCTCCCATTCAAAGCGCTGTTC 420
CY 421 GAVGTACACACCGGAAATCTTCTACGCGCTCCGAGGCGATTAATCTGATGTT 480
CB 421 GAVGTACACACCGGAAATCTTCTACGCGCTCCGAGGCGATTAATCTGATGTT 480
CY 481 AAAGACGCGAGAGAGCTTTGGGTAGATGATGAGAAAGAAAGATGTTCTTCTTCT 540
CB 481 AAAGACGCGAGAGAGCTTTGGGTAGATGATGAGAAAGAAAGATGTTCTTCTTCT 540
CY 541 CTGGAAGTCTCACTGAGAAAGAGATCAATCTTAATGATTGAGAGCAAAATTGAA 600
CB 541 CTGGAAGTCTCACTGAGAAAGAGATCAATCTTAATGATTGAGAGCAAAATTGAA 600
CY 601 GGTAGTATCTGTCGTTGGCGGTGTTGCTCTTAGAGTCTCTCTTGAGATTGACAT 660
CB 601 GGTAGTATCTGTCGTTGGCGGTGTTGCTCTTAGAGTCTCTCTTGAGATTGACAT 660

```

```

CY 661 CTATGTAACATATGCTGTGAGAGAAAC...GCTGTGTTCTTGATTCGATTTGG 720
CB 661 CTATGTAACATATGCTGTGAGAGATCTTGTGTGTGTTCTGATTTGCTGTTGG 720
CY 721 ATCTGATCGTTTGATTCATTTACCTTACGTAACAAATTAATGAAATTAATCGGGA 780
CB 721 ATCTGATCGTTTGATTCATTTACCTTACGTAACAAATTAATGAAATTAATCGGGA 780
CY 781 TTTCCTGTT 789
CB 781 TTTCCTGTT 789

RESULT 2
AAC35978
ID AAC35978 standard; DNA; 522 BP.
XX
AC AAC35978;
XX
ET 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 12108.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EP1031405-A2.
XX
PD 06-SEP-2000.
XX
CY 25-FEB-2000; 200CEP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 08-APR-1999; 99US-0128234.
XX
PR 15-APR-1999; 99US-0128714.
XX
PR 19-APR-1999; 99US-0129845.
XX
PR 21-APR-1999; 99US-0130077.
XX
PR 23-APR-1999; 99US-0130449.
XX
PR 28-APR-1999; 99US-0130510.
XX
PR 30-APR-1999; 99US-0130891.
XX
PR 03-MAY-1999; 99US-0131449.
XX
PR 06-MAY-1999; 99US-0132048.
XX
PR 10-MAY-1999; 99US-0132814.
XX
PR 13-MAY-1999; 99US-0133407.
XX
PR 05-MAY-1999; 99US-0133484.
XX
PR 06-MAY-1999; 99US-0133485.
XX
PR 06-MAY-1999; 99US-0134486.
XX
PR 07-MAY-1999; 99US-0134487.
XX
PR 11-MAY-1999; 99US-0134863.
XX
PR 14-MAY-1999; 99US-0134256.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.
XX
PR 18-MAY-1999; 99US-0134376.
XX
PR 19-MAY-1999; 99US-0134541.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135353.
XX
PR 24-MAY-1999; 99US-0135625.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 27-MAY-1999; 99US-0136382.
XX
PR 28-MAY-1999; 99US-0136782.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137525.
XX
PR 04-JUN-1999; 99US-0137502.

```

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143562.
PR 14-JUL-1999; 99US-0143664.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148177.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0148420.
PR 20-AUG-1999; 99US-0148722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.0%; Score 489.2; DB 21; Length 522;
Best Local Similarity 99.0%; Pred. No. 3e-16;
Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

PD 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123160.
 ER 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128274.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134370.
 PR 16-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136382.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 06-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 16-JUN-1999; 99US-0139457.
 PR 16-JUN-1999; 99US-0139458.
 PR 16-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 16-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139889.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140685.
 PR 25-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141892.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145152.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 23-JUL-1999; 99US-0145275.
 PR 27-JUL-1999; 99US-0145912.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147152.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147491.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149724.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151086.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155135.

```
PR 23-SEP-1999; 99US-0155486-
PR 24-SEP-1999; 99US-0155659-
PR 28-SEP-1999; 99US-0156456-
PR 29-SEP-1999; 99US-0156586-
PR 04-OCT-1999; 99US-0157117-
PR 05-OCT-1999; 99US-0157753-
PR 06-OCT-1999; 99US-0157865-
PR 07-OCT-1999; 99US-0158029-
PR 08-OCT-1999; 99US-0158232-
PR 12-OCT-1999; 99US-0158368-
PR 13-OCT-1999; 99US-0159293-
PR 13-OCT-1999; 99US-0159294-
PR 13-OCT-1999; 99US-0159295-
PR 14-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159330-
PR 14-OCT-1999; 99US-0159331-
PR 14-OCT-1999; 99US-0159637-
PR 14-OCT-1999; 99US-0159638-
PR 18-OCT-1999; 99US-0159584-
PR 21-OCT-1999; 99US-0160741-
PR 21-OCT-1999; 99US-0160767-
PR 21-OCT-1999; 99US-0160768-
PR 21-OCT-1999; 99US-0160770-
PR 21-OCT-1999; 99US-0160814-
PR 21-OCT-1999; 99US-0160815-
PR 22-OCT-1999; 99US-0160980-
PR 22-OCT-1999; 99US-0160981-
PR 22-OCT-1999; 99US-0160989-
PR 25-OCT-1999; 99US-0161404-
PR 25-OCT-1999; 99US-0161405-
PR 25-OCT-1999; 99US-0161406-
PR 26-OCT-1999; 99US-0161359-
PR 26-OCT-1999; 99US-0161360-
PR 26-OCT-1999; 99US-0161361-
PR 28-OCT-1999; 99US-0161920-
PR 28-OCT-1999; 99US-0161992-
PR 28-OCT-1999; 99US-0161993-
PR 29-OCT-1999; 99US-0162142-
Query Match 35.8%; Score 282.8; DB 21; Length 477;
Best Local Similarity 99.3%; Pred. No. 2,1e-74;
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATCATCAACAAAACATTCGATACACAAAACCAACAAAGATTATTTCT 60
DB 3 ATCATCAACAAAACATTCGATACACAAAACCAACAAAGATTATTTCT 62
QY 61 TGAAGAAAGATGAGTTCTCAAGCAAGAGATGACAGTGGAGATGAGCCGTA 120
DB 63 TGAAGAAAGATGAGTTCTCAAGCAAGAGATGACAGTGGAGATGAGCCGTA 122
QY 121 GAGGATTAAGAACCACTAGATCTTTGTCGAGTACATACCTCGGTGGTTAT 180
DB 123 GAGGATTAAGAACCACTAGATCTTTGTCGAGTACATACCTCGGTGGTTAT 182
QY 181 CAACATCTCCGAAACACGTTAGATCTGTTTCTCAAGGAAAGATTCTCTT 240
DB 183 CAACATCTCCGAAACACGTTAGATCTGTTTCTCAAGGAAAGATTCTCTT 242
QY 241 GTCTCCGAGCGGTTACCTCTCTGTTAGAGCGGAGAGAGAGA 286
DB 243 GTCTCCGAGCGGTTACCTCTCTGTTAGAGCGGAGAGAGAGAGA 288
RESULT 5
ID AAC36826 standard; DNA; 413 BP.
XX AAC36826;
AC AAC36826;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15206.
```

```
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
FN EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825-
PR 05-MAR-1999; 99US-0123180-
PR 09-MAR-1999; 99US-0123548-
PR 23-MAR-1999; 99US-0125788-
PR 25-MAR-1999; 99US-0126264-
PR 29-MAR-1999; 99US-0126785-
PR 01-APR-1999; 99US-0127462-
PR 06-APR-1999; 99US-0128234-
PR 08-APR-1999; 99US-0128714-
PR 16-APR-1999; 99US-0129845-
PR 19-APR-1999; 99US-0130077-
PR 21-APR-1999; 99US-0130449-
PR 23-APR-1999; 99US-0130510-
PR 25-APR-1999; 99US-0130891-
PR 28-APR-1999; 99US-0131449-
PR 30-APR-1999; 99US-0132048-
PR 04-MAY-1999; 99US-0132407-
PR 04-MAY-1999; 99US-0132484-
PR 05-MAY-1999; 99US-0132485-
PR 06-MAY-1999; 99US-0132486-
PR 07-MAY-1999; 99US-0132487-
PR 11-MAY-1999; 99US-0132863-
PR 14-MAY-1999; 99US-0134256-
PR 14-MAY-1999; 99US-0134218-
PR 14-MAY-1999; 99US-0134219-
PR 14-MAY-1999; 99US-0134320-
PR 15-MAY-1999; 99US-0134370-
PR 18-MAY-1999; 99US-0134768-
PR 19-MAY-1999; 99US-0134941-
PR 20-MAY-1999; 99US-0135124-
PR 21-MAY-1999; 99US-0135353-
PR 24-MAY-1999; 99US-0135629-
PR 25-MAY-1999; 99US-0136021-
PR 27-MAY-1999; 99US-0136392-
PR 28-MAY-1999; 99US-0136782-
PR 01-JUN-1999; 99US-0137222-
PR 03-JUN-1999; 99US-0137528-
PR 04-JUN-1999; 99US-0137502-
PR 07-JUN-1999; 99US-0137724-
PR 08-JUN-1999; 99US-0138094-
PR 10-JUN-1999; 99US-0138640-
PR 10-JUN-1999; 99US-0138847-
PR 14-JUN-1999; 99US-0139119-
PR 16-JUN-1999; 99US-0139452-
PR 16-JUN-1999; 99US-0139453-
PR 17-JUN-1999; 99US-0139492-
PR 18-JUN-1999; 99US-0139455-
PR 18-JUN-1999; 99US-0139456-
PR 18-JUN-1999; 99US-0139457-
PR 18-JUN-1999; 99US-0139458-
PR 18-JUN-1999; 99US-0139459-
PR 18-JUN-1999; 99US-0139460-
PR 18-JUN-1999; 99US-0139461-
PR 18-JUN-1999; 99US-0139462-
PR 18-JUN-1999; 99US-0139463-
PR 18-JUN-1999; 99US-0139750-
PR 18-JUN-1999; 99US-0139763-
PR 21-JUN-1999; 99US-0139817-
PR 22-JUN-1999; 99US-0139899-
```


7:0417
AAC40417 standard; DNA; 437 BP.
AAC40417;
17-OCT-2000 (first entry)
Arabidopsis thaliana DNA fragment SEQ ID NO: 28228.
Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway;
metabolic pathway; promoter; termination sequence; ss.
Arabidopsis thaliana.
EP1033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-0301439.
25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0133160.
09-MAR-1999; 99US-0123548.
23-MAR-1999; 99US-0125788.
25-MAR-1999; 99US-0126264.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.
06-APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.
16-APR-1999; 99US-0129845.
19-APR-1999; 99US-0130077.
21-APR-1999; 99US-0130449.
23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891.
28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0132048.
30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
11-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136352.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138054.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139452.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 25-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148968.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

```
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154739.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 14.6%; Score 115.4; DB 21; Length 437;
Best Local Similarity 77.3%; Pred. No. 2.9e-24;
Matches 140; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 63 AAGAAGATGAGTCTTACAGCAAGCATGACAGCTGGCATGTGCGAGCCGTAGA 122
DB 50 AAGAAGATGAGTCTTACAGCAAGCATGACAGCTGGCATGTGCGAGCCGTAGA 109
QY 123 GGCATTAAAGACCAACTAGTCTTTGCGGTGAGACTACACTACCTCCGATGATTAATCA 182
DB 110 GGCATTAAAGACCAACTAGTCTTTGCGGTGAGACTACACTACCTCCGATGATTAATCA 169
QY 183 ACATCTCCGGAACACAGTTAGTCTGTTTTCAGAGGGAAGGTTCTCTGCTTCTGCT 242
DB 170 GATCTCCGGAACACAGTTAGTCTGTTTTCAGAGGGAAGGTTCTCTGCTTCTGCT 229
```

```
CY 243 C 243
Do 230 C 230

RESULT 7
AAC50186
ID AAC50186 standard; DNA; 455 bp.
AC AAC50186;
XX 18-OCT-2000 (first entry)
CX Arabidopsis thaliana DNA fragment SEQ ID NO: 63895.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
CX
CX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 200CEP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 05-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 23-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 04-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134376.
XX 18-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
```

PR	16-ANG-1599;	99US-0149356
PR	17-ANG-1599;	99US-0149375
PR	18-ANG-1599;	99US-0149376
PR	20-ANG-1599;	99US-0149372
PR	20-ANG-1599;	99US-0149373
PR	20-ANG-1599;	99US-0149392
PR	21-ANG-1599;	99US-0149392
PR	22-ANG-1599;	99US-0149930
PR	23-ANG-1599;	99US-0150066
PR	25-ANG-1599;	99US-0150884
PR	26-ANG-1599;	99US-0150884
PR	27-ANG-1599;	99US-0151065
PR	27-ANG-1599;	99US-0151066
PR	27-ANG-1599;	99US-0151080
PR	30-ANG-1599;	99US-0151030
PR	31-ANG-1599;	99US-0151348
PR	01-SEP-1599;	99US-0151393
PR	07-SEP-1599;	99US-0152383
PR	10-SEP-1599;	99US-0153070
PR	11-SEP-1599;	99US-0153758
PR	15-SEP-1599;	99US-0154039
PR	16-SEP-1599;	99US-0154039
PR	20-SEP-1599;	99US-0155479
PR	22-SEP-1599;	99US-0155439
PR	23-SEP-1599;	99US-0155466
PR	24-SEP-1599;	99US-0155659
PR	26-SEP-1599;	99US-0156448
PR	29-SEP-1599;	99US-0156566
PR	04-OCT-1599;	99US-0157117
PR	05-OCT-1599;	99US-0157573
PR	06-OCT-1599;	99US-0157865
PR	07-OCT-1599;	99US-0158029
PR	08-OCT-1599;	99US-0158212
PR	12-OCT-1599;	99US-0158369
PR	13-OCT-1599;	99US-0159224
PR	13-OCT-1599;	99US-0159293
PR	13-OCT-1599;	99US-0159293
PR	13-OCT-1599;	99US-0159293
PR	13-OCT-1599;	99US-0159293
PR	21-OCT-1599;	99US-0160717
PR	21-OCT-1599;	99US-0160717
PR	21-OCT-1599;	99US-0160768
PR	21-OCT-1599;	99US-0160774
PR	21-OCT-1599;	99US-0160815
PR	21-OCT-1599;	99US-0160815
PR	22-OCT-1599;	99US-0160999
PR	22-OCT-1599;	99US-0160999
PR	22-OCT-1599;	99US-0161401
PR	25-OCT-1599;	99US-0161405
PR	25-OCT-1599;	99US-0161405
PR	25-OCT-1599;	99US-0161359
PR	26-OCT-1599;	99US-0161359
PR	26-OCT-1599;	99US-0161361
PR	28-OCT-1599;	99US-0161920
PR	28-OCT-1599;	99US-0161920
PR	29-OCT-1599;	99US-0161923
PR	29-OCT-1599;	99US-0162143

Query Match	14.2%	Score 112.2	DB 21	Length 455
Best Local Similarity	76.2%	Pred. No. 2.7e-23		
Matches 138	Conservative 0	Mismatches 43	Indels 0	Gaps 0
QY	63	PLGAAAGTGAAGTTCACAAAGCAACACATGCACATGSCATGAGCATCGGACCCGTAGA	122	
DB	50	AAGCAATGAGCTCTGCACGAACCAACGATGATGCTTCACCAAGCATCGAACCCCTTGA	109	
QY	123	GGCATTAAGAACCAACTAGCTTTTGTCGGTGGAACTACACTACCTCGGTCGTTATACA	182	
DB	110	GGCATGAAACACCACTACGCGCTGTGTGTGGAACTACGATCGGATCTCGAATATA	169	

Query Match 13.7%; Score 108.4; DB 21; Length 252;
 %est Local Similarity 76.4%; Pred. No. 2.3e-22;
 %atches 133; Conservative 0; Mismatches 41; Indels 0; Gaps

70 ATGAGTTCTACAAAGCAGACATGACAGTGGCACTGACATCGAGACCTAAGGACATTA 129

Db	1	ATAGAGTCTGCAGAGCAAAAGCTGATGTTGCTGCAGAGCATGGAGCCGTTGAGCATTTG	60
Qy	130	AAAGACCACCTAGCTGCTTTGTCGGTGGAACTACATCTCCGTCGGTAATCAATCTC	189
Db	61	AAAGACCAACTGAGGGTGTGTCGTTGGAACTAGTATCGATCTGCCAATCAGTATCTA	120
Qy	190	CGGAACAAGCTAGATCTGTTTCTCAAGGGAAGAGTTCTTTCGCTTCTTGTC	243
Db	121	CGGAACAACATTAGATCCGTGTCGCAAGCTAGAGACTCTTCTCTATCATTC	174
RESULT 9			
AAC44432			
ID	AAC44432	standard; DNA; 506 bp.	
XX	AC	AAC44432;	
XX	DT	18-OCT-2000 (first entry)	
XX	XX		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 42806.		
XX	XX		
KM	Hybridization assay; genetic mapping; gene expression control;		
KM	protein identification; signal transduction pathway;		
KM	metabolic pathway; promote; termination sequence; ss.		
XX	OS	Arabidopsis thaliana.	
XX	FN	EP1033405-A2.	
XX	ED	06-SEP-2000.	
XX	FF	25-FEB-2000; 2000EP-0301439.	
XX	XX		
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	21-MAR-1999;	99US-0125789.	
PR	23-MAR-1999;	99US-0128264.	
PR	29-MAR-1999;	99US-0126789.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0128845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0134449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135355.	
PR	24-MAY-1999;	99US-0135625.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	

DB 402 CCGCAGAAACATGTCAGCTTTAAAGAAAAGATTCACCGCAGACGACT 461
 QY 357 AAGCAATACAGCGCAGCGAATCAAAAGCCGATCTACGTCG 401
 DB 462 AAGCAATACAGCGCAGCGAATCAAAAGCGATCTACGTCG 506
 RESULT 10
 AA261724
 ID AA261724 standard; cDNA; 555 BP.
 AC AA261724;
 XX
 XX 27-PAR-2000 (first entry)
 DE cDNA encoding rat dermal papilla protein Dp3, SEQ ID NO:119.
 XX
 XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KM embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KM secreted; transmembrane; inflammation; cancer; neurological disease;
 KM angiogenesis; tumour vascularisation; growth disorder;
 KM developmental disorder; skin wound; hair follicle disorder;
 KM anti-inflammatory; cyostatic; neuroprotective; vulnery; ss.
 XX
 OS Rattus sp.
 XX
 XX MO995865-A1.
 PN
 PD 04-NOV-1999.
 XX
 XX 29-APR-1999; 99MO-NZ00051.
 PP
 XX 29-APR-1998; 98US-0069726.
 PR 09-NOV-1998; 98US-0168930.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA
 PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 XX
 XX WPI: 2000-072177/06.
 DR P-PSDB; AAY76019.
 XX
 XX Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer -
 XX
 XX Claim 1: Page 100; 235pp; English.
 XX
 XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC modulate skin inflammation, to modulate epithelial cell growth and to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AA261606-261832 represent cDNA sequences derived
 CC from several mouse, rat or human skin cell types. Sequences
 CC AA261606-261649, AA261725-261765, AA261802-261811 and AA261826 encode
 CC proteins with an N-terminal signal sequence, indicating that the proteins
 CC are secreted. Sequences AA261650-261668, AA261766-261780, AA261812-261817
 CC and AA261827-261829 encode proteins with one or more putative
 CC transmembrane domains.
 XX
 SQ Sequence 655 BP; 155 A; 180 C; 203 G; 117 T; 0 other;
 Query Match 11.6%; Score 91.2; DB 21; Length 555;
 Best Local Similarity 60.7%; Pred. No. 6.4e-17;
 Matches 147; Conservative 1; Mismatches 94; Indels 0; Gaps 0;
 QY 340 TTCACCGCAGACGACTAAGCAATACAGCGCAGCGAATCAAAAGCGATCTACGTC 339

DB 136 TTCACCGCAGACGACTGCCCCCTACACCGGAGAGGATCAACCACTTACTTG 195
 QY 400 GCAATCAAAAGCCGCTGTTGAYGTACCAACCGAAATCTTTACGCTCCGAGGC 459
 DB 196 GCAGTGAAGGAGGTGTGTGATGTCTACCTCTGGGAAGAGTTTATGACGTGAGGC 255
 QY 460 GATTACTGATGTTCCCGGAAAGACCGGACGAGACTTTGGGTAAATGATGAAGC 519
 DB 256 CCTTACAAAGCCCTTGGCCGGAGAGACTCGACGAGAGTGTGGCAAGATGTGCTGAT 315
 QY 520 GAGAGATGTGTCTCTCTTTGAAGTCTCACTGAGAAAGATCAATCTTTAT 579
 DB 316 CCTCAGACCTCAGTCAATGATTTCTGTCTCACTGCCAAGAGCTGGAAACCTCGAT 375
 QY 580 GA 581
 DB 376 GA 377
 RESULT 11
 AAC99657
 ID AAC99657 standard; cDNA; 655 BP.
 AC AAC99657;
 XX
 XX 08-PAR-2001 (first entry)
 DE Skin cell cDNA, SEQ ID NO: 119.
 XX
 XX Rat; skin cell; cyostatic; anti-inflammatory; anti-HIV;
 KM nocretic; neuroprotective; vulnery; immunomodulatory; vaccine;
 KM keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KM inflammation; neurological disease; ss.
 XX
 OS Rattus sp.
 XX
 XX MO200069294-A2.
 PN
 PD 23-NOV-2000.
 XX
 XX 15-MAR-2000; 2000MO-NZ00075.
 PP
 XX 14-MAY-1999; 99US-0312283.
 PR
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA
 PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble A, Murison JG;
 XX
 XX WPI: 2001-007495/01.
 DR P-PSDB; AAB55958.
 XX
 XX New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 PT disease, cancer and neurological diseases -
 XX
 XX Claim 1: Page 133-134; 352pp; English.
 XX
 XX The present polynucleotide encodes a polypeptide which is expressed in
 CC mammalian skin cells. The polypeptide is useful for stimulating
 CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
 CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
 CC tumours, modulating skin inflammation, stimulating the growth of
 CC epithelial cells, inhibiting the binding of human immunodeficiency virus
 CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
 CC neurological diseases. The polynucleotide can be used as a marker, in
 CC the identification of genetic disorders, and for the design of
 CC oligonucleotides for examining expression patterns.
 XX
 SQ Sequence 655 BP; 155 A; 180 C; 203 G; 117 T; 0 other;
 Query Match 11.6%; Score 91.2; DB 22; Length 655;
 Best Local Similarity 60.7%; Pred. No. 6.4e-17;

Matches 147; Conservative 1; Mismatches 94; Indels 0; Gaps 0;

```

QY 340 TTCACCGGAGGAGCTAAGCCATACAGCGACCGAGCATCAACCCGCTACGTC 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 TTCACCGGAGGAGAGCTGGCCCGCTACAGCGGCGAGAGAGATCAACCCGCTACTTG 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 400 GCAATCAAAAGCCCGTGTGTTGATGATCCACCGGAAAATCCTTACCGCTCCGAGGC 459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 GCAGTGAAGGAGAGGTGTTGATGATCACTCGGGAAGAGTTTATGAGAGTGGAGCC 255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 460 GATTCTCGATGTTGCCCGGAAAAGCGGACGAGAGCTTTGGGTAAAGTGAAGAAC 519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CCCTACAAAGCGCTTGCGCGGAAAGACTCGAGCAGAGGTGTGCGCAAGTGTGCTGGAT 315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 520 GAAGAAGATGTGTCTCTCTCTTGAAGCTCTCAGTGAAGAAAGATCAATCTTTAAT 579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 CCTGCAAGCTCAGTCAATGATTTGTGTCTCACTGCACTGCCAAGAGACTGGAAGCCTCGAT 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 580 GA 581
    ||
Db 376 GA 377
    ||

```

RESULT 12

ABL34809
ID ABL34809 standard; cDNA; 655 BP.

AC ABL34809;

DT 04-APR-2002 (first entry)

DE Rat cDNA isolated from skin cells SEQ ID NO: 119.

DE Human; rat; mouse; skin cell; skin wound; cancer; growth defect;

KM developmental defect; inflammatory disease; dermatological; vulnereary;

KM immunomodulator; anti-inflammatory; cyostatic; neuroprotective; gene;

OS Rattus sp.

PN W0200190357-A1.

PD 29-NOV-2001.

PF 24-MAY-2001; 2001WO-N200099.

PR 24-MAY-2000; 2000US-206650P.

PR 25-JUL-2000; 2000US-221232P.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Rumbi, KD;

DR WPI, 2002-122020/16.

PT New polynucleotides and polypeptides encoded by the polynucleotides;

PT isolated from skin cells, useful for treating skin wounds, cancers,

PT growth and developmental defects, inflammatory diseases, or for

PT modulating immune responses

PS Claim 1; Page 116; 466pp; English.

CC The present invention provides the protein and coding sequences of cDNAs

CC isolated from human, murine and rat skin cell libraries. The sequences

CC can be used in the development of therapeutic agents useful in the

CC treatment of skin diseases, including skin wounds, cancer, growth

CC defects, developmental defects and inflammatory diseases. The proteins

CC have important roles in the induction of hair growth, cell proliferation

CC and cell-cell interaction, in maintaining tissue integrity, in wound

CC healing and in modulating immune responses. The present sequence is a

CC cDNA of the invention.

CC Sequence 655 BP; 155 A; 180 C; 203 G; 117 T; 0 Other;

Query Match 11.6%; Score 91.2; DB 24; Length 655;

Best Local Similarity 60.7%; Pred. No. 6.4e-17;

Matches 147; Conservative 1; Mismatches 94; Indels 0; Gaps

```

QY 340 TTCACCGGAGGAGCTAAGCCATACAGCGACCGAGCATCAACCCGCTACGTC 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 TTCACCGGAGGAGAGCTGGCCCGCTACAGCGGCGAGAGAGATCAACCCGCTACTTG 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 400 GCAATCAAAAGCCCGTGTGTTGATGATCCACCGGAAAATCCTTACCGCTCCGAGGC 459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 GCAGTGAAGGAGAGGTGTTGATGATCACTCGGGAAGAGTTTATGAGAGTGGAGCC 255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 460 GATTCTCGATGTTGCCCGGAAAAGCGGACGAGAGCTTTGGGTAAAGTGAAGAAC 519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CCCTACAAAGCGCTTGCGCGGAAAGACTCGAGCAGAGGTGTGCGCAAGTGTGCTGGAT 315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 520 GAAGAAGATGTGTCTCTCTCTTGAAGCTCTCAGTGAAGAAAGATCAATCTTTAAT 579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 CCTGCAAGCTCAGTCAATGATTTGTGTCTCACTGCACTGCCAAGAGACTGGAAGCCTCGAT 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 580 GA 581
    ||
Db 376 GA 377
    ||

```

RESULT 13

AEQ65424
ID AEQ65424 standard; DNA; 751 BP.

AC AEQ65424;

DT 21-AUG-2002 (first entry)

DE Arabidopsis thaliana polynucleotide SEQ ID NO 1.

KM Arabidopsis thaliana, chole cress; plant; transgenic; GMO; disease;

KM stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;

KM insecticide; antibiotic; ds.

OS Arabidopsis thaliana.

PN US2002059663-A1.

PD 16-MAY-2002.

PF 26-JAN-2001; 2001US-0770149.

PR 27-JAN-2000; 2000US-178506P.

PA (GURL/) GORLACH J.

PI (KATY/) AN Y.

PI (HAM/) HAMILTON C M.

PI (PRICE/) PRICE J L.

PI (RAINE/) RAINE T M.

PI (YU/) YU Y.

PI (PAGE/) PAGE A.

PI (MATH/) MATH A V.

PI (LEDF/) LEDFORD B L.

PI (MOES/) MOESSNER J P.

PI (HAAS/) HAAS W D.

PI (GARC/) GARCIA C A.

PI (KRIK/) KRICKER M.

PI (SLAT/) SLATER T.

PI (DAVI/) DAVIS K R.

PI (ALLE/) ALLEN K.

PI (HOFF/) HOFFMAN N.

PI (HURB/) HURBAN P.

PI Gortach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

PI Ramesha JG, Page A, Mathew AV, Ledford BL, Moessner JP, Haas WD;

PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

PI Human P:
 XX WPI: 200, 479224/51.
 XX
 PT New nucleic acid that hybridizes to Arabidopsis thaliana sequences,
 PT useful e.g. for preparing transgenic plants with increased resistance
 PT or altered metabolism
 XX
 PS Claim 1, SEQ ID NO 1, 40pp + Sequence Listing; English.
 XX
 CC The invention relates to nucleic acids (I) that hybridize under stringent
 CC conditions to any of 999 sequences (AB065424-AB066422) or their
 CC fragments. (II) are used to express the corresponding polypeptides (II) or
 CC to produce genetically modified plant cells or transgenic plants, which
 CC may have improved resistance to disease or stress, or altered
 CC metabolic/biosynthetic pathways (for production of commercial),
 CC nutritional or medicinal products), or generally any trait of interest,
 CC or can be used to screen for biologically active agents (e.g. fungicides,
 CC insecticides and antibiotics).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO at seqdata.uspto.gov/sequence.html?docid=999909770142.
 CC
 SQ Sequence 751 BP, 195 A, 164 C, 183 G, 209 T, 0 other;
 Query Match 11.4%; Score 89.6; DB 24; Length 751;
 Best Local Similarity 56.6%; Pred. No. 2,1e-16;
 Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;
 QY 337 GAATTCACCGCAGGAGCTAAGCCATACAGCGCAGCGAATCAAGCCGCTAC 396
 DB 186 GAGATCACGAGGAGGAGCTTAACAGTACGATGCTGATCTCTCAAAAGCCCTTCTT 245
 QY 337 GTCCATCAAGAGCCGCTGCTGTCAGCAGCCGAAATCCCTTACGCGCCCGGA 456
 DB 246 ATGCTATCAACATGATCATATGATATGATACAAACAGAGATGTTACGCGACGGA 305
 QY 457 GGCATTAATCGATGTCGCGGAAAGAGCGGAGCAGAGCTTTGGGTAAATAGTAG 516
 DB 306 GGAACATATGCTTTGTTGACGAGAAAGAGCGTACGCGCTCTTGCAAGATATGCTTT 365
 QY 517 AACAGAGAGATGTCCTCTCTTGAAGTCTCTACTGAGAAAGATCAATCTTT 576
 DB 366 GAGCAGAAAGACTGATGCGAGTGTCTGCTGCTCTGCTCTTGAGCTAGATGCTCTT 425
 QY 577 AATATTGGAGAGCAATTTGAAGCTATGCTGTTGGCGCGT 626
 DB 426 CAGATTTGGAGTACAGTTTCATAGCAAGTATGCTTGAAGTGTGACTGT 475
 RESULT 14
 AAC50344
 ID AAC50344 standard; DNA: 887 BP.
 AC AAC50344:
 XX 1P-OCT-2700 (first entry)
 DT
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 64487.
 XX
 XX Hybridisation assay; genetic mapping; gene expression control;
 KM Protein identification; signal transduction pathway;
 KM Metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 XX E71033405-A2.
 PN
 XX
 XX 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-0201439.
 PP
 XX 25-FEB-1999; 99US-0121825.
 PR

PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125768.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 09-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139464.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140685.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.

Arabidopsis thaliana.
EP1033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-030139.
25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.
23-MAR-1999; 99US-0125788.
25-MAR-1999; 99US-0126264.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.
06-APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.
16-APR-1999; 99US-0129845.
19-APR-1999; 99US-0130077.
21-APR-1999; 99US-0130449.
23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891.
28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0132048.
30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
11-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134321.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
18-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
ER 01-JUL-1999; 99US-0141842.
ER 01-JUL-1999; 99US-0142154.
ER 02-JUL-1999; 99US-0142055.
ER 06-JUL-1999; 99US-0142390.
ER 08-JUL-1999; 99US-0142803.
ER 09-JUL-1999; 99US-0142920.
ER 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143622.
ER 15-JUL-1999; 99US-0144005.
ER 16-JUL-1999; 99US-0144085.
ER 19-JUL-1999; 99US-0144325.
ER 19-JUL-1999; 99US-0144331.
ER 19-JUL-1999; 99US-0144332.
ER 19-JUL-1999; 99US-0144333.
ER 19-JUL-1999; 99US-0144334.
ER 19-JUL-1999; 99US-0144335.
ER 20-JUL-1999; 99US-0144632.
ER 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
ER 21-JUL-1999; 99US-0145086.
ER 21-JUL-1999; 99US-0145088.
ER 22-JUL-1999; 99US-0145087.
ER 22-JUL-1999; 99US-0145089.
ER 22-JUL-1999; 99US-0145192.
ER 23-JUL-1999; 99US-0145145.
ER 23-JUL-1999; 99US-0145218.
ER 23-JUL-1999; 99US-0145222.
ER 26-JUL-1999; 99US-0145276.
ER 27-JUL-1999; 99US-01452913.
ER 27-JUL-1999; 99US-01452919.
ER 28-JUL-1999; 99US-0145591.
ER 02-AUG-1999; 99US-0146366.
ER 02-AUG-1999; 99US-0146388.
ER 02-AUG-1999; 99US-0146389.
ER 03-AUG-1999; 99US-0147038.
ER 04-AUG-1999; 99US-0147204.
ER 04-AUG-1999; 99US-0147302.
ER 05-AUG-1999; 99US-0147319.
ER 05-AUG-1999; 99US-0147320.
ER 06-AUG-1999; 99US-0147320.
ER 06-AUG-1999; 99US-0147303.
ER 09-AUG-1999; 99US-0147416.
ER 09-AUG-1999; 99US-0147493.
ER 10-AUG-1999; 99US-0147935.
ER 11-AUG-1999; 99US-0148171.
ER 11-AUG-1999; 99US-0148319.
ER 12-AUG-1999; 99US-0148319.
ER 13-AUG-1999; 99US-0148321.
ER 13-AUG-1999; 99US-0148365.
ER 13-AUG-1999; 99US-0148684.
ER 16-AUG-1999; 99US-0149368.
ER 17-AUG-1999; 99US-0149175.
ER 18-AUG-1999; 99US-0149426.
ER 20-AUG-1999; 99US-0149722.
ER 20-AUG-1999; 99US-0149723.
ER 20-AUG-1999; 99US-0149723.
ER 23-AUG-1999; 99US-0149929.
ER 23-AUG-1999; 99US-0149902.
ER 23-AUG-1999; 99US-0149930.
ER 25-AUG-1999; 99US-0150566.
ER 26-AUG-1999; 99US-0150884.
ER 27-AUG-1999; 99US-0151065.
ER 27-AUG-1999; 99US-0151066.
ER 27-AUG-1999; 99US-0151080.
ER 30-AUG-1999; 99US-0151080.
ER 31-AUG-1999; 99US-0151303.
ER 31-AUG-1999; 99US-0151438.
ER 01-SEP-1999; 99US-0151930.
ER 07-SEP-1999; 99US-0152363.
ER 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

```
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

```
Query Match 11.4%; Score 89.6; DB 21; Length 889;
Best Local Similarity 56.6%; Pred. No. 2.2e-16;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;
```

```
QY 337 GAATTACCGGAGAGCAGTAAAGCAATACAGGACCGGATCAAGCCGATCTTAC 396
DB 264 GAGATCAGGAGGAGGAGCTTAAACAGTACGATGCTGTGATCTCAAAAGCCCTTCTT 323
QY 397 GTCCAAATCAAAGGCGGCTGTCGATGTCACCCGAAATCCTTCTACGGCTCCGGA 456
DB 324 ATGCTATCAACATCATGATCTATGATGTTACAAAGCAGATGTTCTACGGACACGA 383
QY 457 GGCATTACTCGATGTCGCGCGGAAAGAGCGAGAGCTTTGGCTAAGTACATCTT 516
DB 384 GGACCATATGCTTTGTTGAGGAAAGAGCTGATCTGTGCTTGTCCCTTGAGCTAATCTCTT 443
QY 517 AACGAGAAGATGTCCTCTCTTGAAGTCTCAGTGAAGAGATCAATCTTCTT 576
DB 444 GAGGAGAAAGCTTGATGATGATCTGTGCTTGTGCTTGTGAGCTAATCTCTT 503
QY 577 AATGATTGGAGGACCAATTTGAAGCTAAGTATCTGTGCTTGGCCGCTGT 626
DB 504 CAAGATTGGAGTACAGTTCATGACCAAGTATGATGATGATGATGATGATGAT 553
```

Search completed: January 8, 2003, 14:09:48
Job time : 271 secs

Search completed: January 8, 2003, 14:12:20
 Job time : 93 secs


```

: Sequence 3, Application US/09565808
: Patent No. 6432674
: GENERAL INFORMATION:
: APPLICANT: Hirata, Yuichi
: TITLE OF INVENTION: STEROID HORMONE BINDING PROTEIN
: FILE REFERENCE: 06501-059001
: CURRENT APPLICATION NUMBER: US/09/565,808
: CURRENT FILING DATE: 2000-05-05
: PRIOR APPLICATION NUMBER: WO/JP98/05010
: PRIOR FILING DATE: 1998-11-06
: PRIOR APPLICATION NUMBER: JP/9/322376
: PRIOR FILING DATE: 1997-11-07
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 672
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(669)
: US-09-565-808-3

```

```

Query Match          7.9%; Score 62.2; DB 4; Length 672;
Best Local Similarity 54.6%; Pred. No. 7e-10;
Matches 171; Conservative 1; Mismatches 129; Indels 12; Gaps 2;

```

```

QY 320 CTTTAAAGAAAAGATGATTCACCGAGACGCTAACGCAATCAACGCGACGAGC 379
DB 287 CTGGATATAGAGAGCGGACTTCAGCTTGAGACGCTTGGCCGAGTACGAGCGCTCCGCA 346
QY 380 AATCAAGCCGATCTACGTCGCAATCAAGCGCGCTGTGTGATGATGACCGCGGAAAT 439
DB 347 ACCC---GCGCATCTGCTCGCGGTCAATGGAAGATCTTGACGTGACCAAGGACAGA 403
QY 440 CCTTAGAGGCTCCGAGGCGCATTAATGATGTCGCGGAAAGACGAGAGAGCTT 499
DB 404 AGCTTACGCGCGCGGCTCATATGGAATATTTCTGTAGGAGATGCTCCAGAGAC 463
QY 500 TGGGTAGATAGTAGAGACGAGAA-----AGATGTCTCTCTCTCTTGTGAAGTGC 550
DB 464 TGGCCACATTTTGGCTAGATTAAGATGACCTTAGAGATGATGATGATCTCTCAGAT 523
QY 551 TCAGTACAGAAAGATCAATACCTTAATGATTTGGAGACCAATTGAAGCTAGATATC 610
DB 524 TGAATGCGATCAATAGAGAGTGTGAGATGCGAAATGACAGTTAAAGAAATATG 583
QY 611 CTGTCTTGGCGC 623
DB 584 ATATGTAAGCAG 596

```

RESULT 3

```

: US-08-960-022-5
: Sequence 5, Application US/08960022
: Patent No. 5976837
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavalie, Edward R.
: APPLICANT: Racie, Lisa A.
: APPLICANT: Merberg, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Spaulding, Vikki
: APPLICANT: Agostino, Michael J.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
: TITLE OF INVENTION: ENCODING THEM
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESS: Genetics Institute, Inc.
: STREET: 87 Cambridgepark Drive
: CITY: Cambridge
: STATE: MA

```

```

: COUNTRY: U.S.A.
: ZIP: 02140
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: CREATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/660,022
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Sprunger, Suzanne A.
: REGISTRATION NUMBER: 41,323
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 498-9284
: TELEFAX: (617) 876-5851
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1868 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-960-022-5

```

```

Query Match          7.0%; Score 55.2; DB 2; Length 1868;
Best Local Similarity 53.7%; Pred. No. 2e-07;
Matches 158; Conservative 3; Mismatches 121; Indels 12; Gaps 2;

```

```

QY 337 GAATTCACCGCAGAGAGTAGCCATATACAGCGCACAGCATCAAGCCGATCTAC 396
DB 265 GACTTACCCCGCGGAGAGTGGGCGCTTCAGCGCGCTCCAGAGCC---GCCATATCT 321
QY 397 GTCCCAATCAAGCGCGGTGTGTGATGATCAACCGGAAATCTTTAGGCTTCGGA 456
DB 322 ATGCCCATACAGCGCAAGGTGTTCATGTGACCAAGCGCGCAAAATTTCTAGSGCCGAG 381
QY 457 GCGCATTTACGATGTGCGCGAAGAACGCGGAGAGCTTTGGGTAGATGATGATAG 516
DB 382 GCGCGCATATGCGGCTCTTCTGTGAGAGATGATCAAGGAGCCTTGGCACATTTTCCCTG 441
QY 517 AACGAGAA-----GATGTCTCTCTCTCTTGAAGGCTCACTAGAAAGATC 567
DB 442 GATAGAGAGCAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
QY 568 AATACCTTTAATGATTTGGAGACCAATTTGAAGCTAAGTATCTGTGTGCTC 621
DB 502 GAGACTGTAGTACTGGAGAGTCTGATTCATTCAAGTATCATCATCAGCTGGGC 555

```

RESULT 4

```

: US-09-565-808-1
: Sequence 3, Application US/09565808
: Patent No. 6432674
: GENERAL INFORMATION:
: APPLICANT: Hirata, Yuichi
: TITLE OF INVENTION: STEROID HORMONE BINDING PROTEIN
: FILE REFERENCE: 06501-059001
: CURRENT APPLICATION NUMBER: US/09/565,808
: CURRENT FILING DATE: 2000-05-05
: PRIOR APPLICATION NUMBER: WO/JP98/05010
: PRIOR FILING DATE: 1998-11-06
: PRIOR APPLICATION NUMBER: JP/9/322376
: PRIOR FILING DATE: 1997-11-07
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 588
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS

```

us-09-649-866a-1.rn1

Page 3

IMMEDIATE SOURCE:

```

Best Local Similarity 4.8%; Score 39.2; DB 1; Length 7219
Matches 7; Conservative 100

```

```

RESULT 7
US-09-462-136-1/c
; Sequence 1, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Caretes, et al.
; TITLE OF INVENTION: Genes For Niemann-Pick Type C Diseases
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3837)
US-09-462-136-1

```

```

Query Match      4.68; Score 36.2; 25 4; Length 453;
Best Local Similarity 48.3%; Pred. No 0.34;
Matches 101; Conservative 0; Mismatches 108; Indels 3; Gaps 0

Qy 11 AAGCAATTCATACACAAAACACAAAACACAAAGAGATTATTTCTGAGAGAGA 70
Db 864 AAAAAATCTCCAAAACACACAGCAAAAACGCCATGTAGTGATCCACATGATATACAT 70
Qy 71 TGAGTTCTACAGCAAAAGCATGGACAGTCGACAGTGAATGAGAGCCGTAGAGATTATA 805
Db 804 CATGGCTCCAAAGCCAAAGATCGTCCAGGAGACAGGAGAGAGAGTGGGGCTTGGGCTTGG 130
Qy 131 AAGCAACTAGAGTCTTGTGCGTGAACACTACATACCTCGGTATATCAACATCTTCC 745
Db 131 AAGCAACTAGAGTCTTGTGCGTGAACACTACATACCTCGGTATATCAACATCTTCC 190

```

Db 744: GCCACGACATATAGCAGCTCTGGCGCTACATGGTGGCTGTGACTTCATCCAGAGATC 689
 Cy 191 GGAACACCTTATGATCTCTTTCTCTAAAGG 219
 Db 684 GTCCACAGCCTTTGGTGGCATGTGTCATGG 656

RESULTS
US-09-149-476-24/C
Sequence 24, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION
APPLICANT: Rosen et al
TITLE OF INVENTION: 166 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: 166 Human Secreted proteins
CURRENT FILING DATE: US/09/149, 476
EARLIER APPLICATION NUMBER: 1998-09-08
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,526
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23

1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 26

EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,892
 EARLIER FILING DATE: 1997-05-22
 EARLIER APPLICATION NUMBER: 60/057,751
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/047,555
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,559
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,558
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,585
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,586
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,590
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,594
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047,559
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/043,578
 EARLIER FILING DATE: 1997-04-11
 EARLIER APPLICATION NUMBER: 60/043,576
 EARLIER FILING DATE: 1997-04-11
 EARLIER APPLICATION NUMBER: 60/047,501
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/043,670
 EARLIER FILING DATE: 1997-04-11
 EARLIER APPLICATION NUMBER: 60/056,632
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,664
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,976
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,831
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,959
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,875
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,862
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,887
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056,998
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/053,964
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/057,650
 EARLIER FILING DATE: 1997-09-05
 EARLIER APPLICATION NUMBER: 60/056,884
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/057,555
 EARLIER FILING DATE: 1997-09-05
 EARLIER APPLICATION NUMBER: 60/045,510
 EARLIER FILING DATE: 1997-06-13
 EARLIER APPLICATION NUMBER: 60/061,060
 Query March
 Ass: Local Similarity 4.5% Score 35.2; DB 4; len=1
 Matches 76; Conservative

688
226

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 14:09:55 ; Search time 2802 Seconds

(without alignments)
8194.901 Million cell updates/sec

Title: US-09-649-866A-1

Perfect score: 789
1 atcattcaacaaacacatc.....caaacg99gattctgctc 789

Scoring table: IDENTITY NUC
Gapop 10-0, Gapext 1.0

Search: 2054610 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

GenEmbl:
1: gb_ba:
2: gb_hg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pac:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vl:
30: em_hg_hum:
31: em_hg_inv:
32: em_hg_other:
33: em_hg_mus:
34: em_hg_pin:
35: em_hg_rtd:
36: em_hg_mam:
37: em_hg_vrt:
38: em_hg_hum:
39: em_hg_mus:
40: em_hg_mus:
41: em_hg_other:

Pred. No. is the number of results predicted by chance to have a

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489.2	62.0	522	8	AY084294
2	487.2	61.0	103495	8	AC006585
3	487.6	62.0	2836	8	AY050994
4	282.3	35.3	477	8	AY085799
5	282.8	35.3	83922	8	ATP24G24
6	283.8	35.3	99856	8	ATP24G24
7	283.8	35.3	99856	8	ATP24G24
8	125.2	15.9	19936	2	AP005115
9	122	15.5	13158	2	AP004054
10	113.8	14.4	83922	8	ATP24G24
11	113.8	14.4	99856	8	ATP24G24
12	113.8	14.4	99856	8	ATP24G24
13	90.2	11.4	11382	2	AC130811
14	88.6	11.4	847	8	AF133284
15	88.6	11.4	889	8	AF133284
16	84.6	10.7	81609	8	AC027035
17	84.6	10.7	97242	8	AC051630
18	79.2	10.0	861	8	AY046006
19	75.8	9.6	702	6	AX412363
20	75.8	9.6	702	8	AY079104
21	75.8	9.6	724	9	AF173937
22	75.8	9.6	930	8	AF153283
23	75.8	9.6	945	8	AF153283
24	75.8	9.6	953	8	BCC08823
25	74.6	9.5	697	9	AK074431
26	70.6	8.5	10000	2	LMFCHP35_23
27	70	8.5	304	9	AF049672
28	69.2	8.8	306	8	LEP52
29	65.8	8.3	509	8	MSA48337
30	65	8.2	48420	8	AB025603
31	62.2	7.5	1081	6	AX454748
32	62.2	7.5	1081	6	AX454748
33	62.2	7.5	1874	9	HS42020
34	62.2	7.5	1897	9	BC016692
35	61.6	7.8	919	8	AY084353
36	59.2	7.5	63554	8	H0410608
37	59.2	7.5	73187	2	OSIG00033
38	59.2	7.5	159525	8	OSJN00019
39	57.4	7.3	2708	9	AK091741
40	56.4	7.1	1490	3	AY061163
41	56	7.1	1893	4	SSSTERMBP
42	55.2	7.0	1868	6	AR083260
43	54.4	6.9	1866	9	RC043739
44	54.4	6.9	1890	6	AX335570
45	54.4	6.9	1890	6	AX336196

ALIGNMENTS

RESULT 1
AY084294 522 bp mRNA linear PLN 21-JUN-2002
Arabidopsis thaliana clone 10261 mRNA, complete sequence.
AY084294
VERSION
KEYWORDS
SOURCE
ORGANISM
Chromosome
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 522)
Haas, B.J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N.,
Feldmann, R.A., Flavell, R.B., White, O. and Salzberg, S.L.

Full-length messenger RNA sequences greatly improve genome annotation
Genome Biol. (2002) In press
2 (bases 1 to 522)
Broyer, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
Full-length cDNA from *Arabidopsis thaliana*
Unpublished
3 (bases 1 to 522)
Broyer, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc., 3007 Malibu Canyon Road
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genes carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers
1..522
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="10261"
54..356
/codon_start=1
/product="putative steroid binding protein"
/protein_id="AA060885.1"
/db_xref="GI:21536553"
/translation="MEFTAEQLSCNGTDSRPIYAKGVFVDTTGKFFGSGTQY SMFKGKSRPLKMGSKNEEVSPELGLERENILNMDETFEAKYVVERVYS"

AGE UNIT 141 a 100 c 121 g 160 t

Over Match 52.0% Score 489.2; DB 8; Length 522;
Res Local Similarity 99.0% Pred. No. 6e-120;
Mat es 491; Conservative 1; Mismatches 4; Indels 0; Gaps 2;

94 TTCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAGATGATTCACCGAGAGCA 353
|||||
14 TTCTTTCAGAAACAATTGATCAGAGCTTTAAAGAAAAGATGATTCACCGAGAGCA 73
54 GCTTAGCGCATACAGCGGACCGAGCAATCAAGTCGATCTATGCCCATCAAGGCGG 413
|||||
74 GCTAAGCATACAGCGGACCGAGCAATCAAGTCGATCTATGCCCATCAAGGCGG 133
|||||
114 TGTATTGAGATGACACCGGAAATCTCTTACGCGCTCCGAGGCGATTACTGATGTT 473
|||||
134 TGTTTTCATGTCACACCGGAAATCTCTTACGCGCTCCGAGGCGATTACTGATGTT 193
|||||
174 CGCCGAGAAAGACCGGACGAGCTTTGGGTAGATGATAGACGAAAGATGATGTC 513
|||||
194 CGCCGAGAAAGACCGGACGAGCTTTGGGTAGATGATAGACGAAAGATGATGTC 253
|||||
334 TCCTTCTTGAAGGCTGCTGCTGAGAAAGATCATACTCTTAATGATTGGAGACCAA 593
|||||
354 TCCTTCTTGAAGGCTGCTGCTGAGAAAGATCATACTCTTAATGATTGGAGACCAA 313
|||||
394 ATTGAGAGTAAAGTATCTGTCGTTGGCGCTGTGTCTCTTAAGTCTCTTCTAGATT 653
|||||
114 ATTGAGAGTAAAGTATCTGTCGTTGGCGCTGTGTCTCTTAAGTCTCTTCTAGATT 373

QY 654 GCACATATGTAATGTAATGTTGTTGAGGATCTTTGTTGTTGTTTCTGATTTCCG 713
|||||
Db 374 SCACATATGTAATGTAATGTTGTTGAGGATCTTTGTTGTTGTTTCTGATTTCCG 433
|||||
QY 714 TTTTGGATTCGATTCGTTTGCATTCACATTCACATTCACATTCACATTCACATTC 773
|||||
Db 434 TTTTGGATTCGATTCGTTTGCATTCACATTCACATTCACATTCACATTCACATTC 493
|||||
QY 774 TCGGGATTTGCTGTT 789
|||||
Db 494 TCGGGATTTGCTGTT 509
|||||

RESULT 2
AC006585/c 103495 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F27C12 map m1238, complete
DEFINITION sequence.

ACCESSION
AC006585
VERSION AC006585.8 GI:20197873
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 103495)
Lin, X., Kaul, S., Shea, T.P., Fujii, C.Y., Shen, M., Vanden, S.E.,
Barnstead, M.E., Mason, T.M., Bowman, C.L., Rong, C.M.,
Benito, M.-I., Carrera, A.J., Greasy, T.H., Buell, C.R., Town, C.D.,
Hierman, N.C., Fraser, C.M. and Venter, J.C.
Unpublished

REFERENCE
2 (bases 1 to 103495)
Lin, X.
AUTHORS Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

REFERENCE
3 (bases 1 to 103495)
Town, C.D. and Kaul, S.
AUTHORS Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org

COMMENT
On Apr 18, 2002 this sequence version replaced g1:6598610.
CSHL/RWGSC/ABI Arabidopsis Genome Sequencing Consortium.
Information on physical mapping and YAC and BAC library
construction as well as added annotation can be viewed at
http://www.cshl.org/Arabweb/. We used GenScan, Grail, and XZEF for
predicting coding exons and assembling genes. BAC F6P23 maps to
YAC
YUPC1C81.

FEATURES
Source
1..103495
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="m1238"
/clone="F27C12"
/complement(1..10309)
/note="overlap with BAC clone F13D4 (AL031369.1..10306)."

repeat_region
complement(100..136)
/rpt_family="AT_rich"
repeat_region
complement(206..236)
/rpt_family="AT_rich"
1028..3697
/gene="At2g25070"
/note="F27C12.1; supported by full length cDNA:
Ceres.11428"

GENE
join(1028..1197,1624..1742,1826..1889,2061..2169,
2257..2410,2506..2755,2850..2926,3010..3062,3149..3203,
3287..3697)
/gene="At2g25070"

[illegible]

Bb	243	TTAAAGGCTCAGCAAGAAAGACATCATACCTTTAANGATTGGAGACCCGAATTGGAG	302
Oy	602	CTAAGTCATCCGRCGCGCCGTGGCTGTCTCTTAAGTCTCTCTTGAGATGCACATAAG	652
Dd	303	CTAAGTATCCGTCGTGGTGGCGCGTGTCTCTTAAGTCTCTCTTAGATATTCGACTAAG	362
Oy	662	TTATGTACTATTGTGTGTAGAGATCTTTGTGTGTGGTGTTCGATTTTCGATTTGG	722
Dd	363	TTAAGTAACTATCTGTGTAGAGATCTTTGTGTGTGGTGTTCGATTTTCGATTTGG	422
Oy	722	TCTGATCCTTTTGTATACATTAACCATTAAGTACCAGTATCTATGAAATTAATCGGGGAT	781
Dd	423	TCTGATCGTWTGTGACATTAACCATTAAGTACCAGTATCTATGAAATTAATCGGGGAT	482
Oy	782	TTCCGTGT 789 	
Dd	483	TTCCGTGT 450	
RESULT 4			
LOCUS	AY085799	477 bp mRNA linear PIN 25-JUN-2002	
DEFINITION	Arabidopsis thaliana clone 1831 mRNA, complete sequence.		
ACCESSION	AY085799		
VERSION	AY085799.1 GI:21404509		
XEMORDS	FLI CDNA.		
SOURCE	chale cress.		
ORGANISM	Arabidopsis thaliana		
	Seemataceae; Viciales; Streptophyta; Embryophyta; Tracheophyta;		
	Eukaryota; Magnoliophyta; Eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi-		
REFERENCE	1 (bases 1 to 477)		
AUTHORS	Haas,B.J., Volkovskiy,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.		
TITLE	Full-length messenger RNA sequences greatly improve genome annotation		
JOURNAL	Genome Biol. (2002) In press		
REFERENCE	2 (bases 1 to 477)		
AUTHORS	Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
TITLE	Full-length cDNA from Arabidopsis thaliana		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 477)		
AUTHORS	Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,		
COMMENT	Malibu, CA 90265, USA		
	This clone sequence is one of 5,000 Ceres full-length cDNAs made available by TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Generated carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.		
FEATURES			
source	1..477		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
	/clone="1831"		
CDS	72..344		
	/codon_start=1		
	/product="probable wound-induced protein"		
	/protein_id="AAW63015.1"		

```

/ab xref="GI:21553934"
/translation="MSRSKMTVAIVSICGVELKNOGLCMYNTLSVNOHLRNT
RVSQGRKFRSSSVSAVTSSESEKAKKEESLRTVMYLSGMDN"
BASE COUNT      164 a      98 c      86 g      129 t
ORIGIN
Query Match      35.8%  Score 282.8; DB:8; Length 477;
Best Local Similarity 99.3%; Pred No. 9.6e-65;
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0
QY 1 ATCATCAACAAAAACAATTTCTGCATACACAAAACACAAACACAAAGATTATTTCTC 60
Db 3 ATCATCAACAAAAACAATTTCTGCATACACAAAACACAAACACAAAGATTATTTCTC 62
QY 61 TGAAGAAAGATGAGTTCTCAAGCCAAAGATGACAGTGGACGTGACATCGAGCCGTA 120
Db 63 TGAAGAAAGATGAGTTCTCAAGCCAAAGATGACAGTGGACGTGACATCGAGCCGTA 122
QY 121 GAGGCATTAAAGACCAACTAGTCTTTTGTGCGTGAACCTACATCCGGTGGTTAAT 180
Db 123 GAGGCATTAAAGACCAACTAGTCTTTTGTGCGTGAACCTACATCCGGTGGTTAAT 182
QY 181 CAATATCTCCGAGACAGCTTACATCTGTTTCTCAAGGAAAAAGTTCTCTTCTTCT 240
Db 183 CAATATCTCCGAGACAGCTTACATCTGTTTCTCAAGGAAAAAGTTCTCTTCTTCT 242
QY 241 GTCTCCGACATCGTTACCTCTCTCTGTGTAGAGCGGAAAGCAAGAA 286
Db 243 GTCTCCGACATCGTTACCTCTCTCTGTGTAGAGCGGAAAGCAAGAA 288
RESULT 5
LOCUS      T9A4
DEFINITION Arabidopsis thaliana BAC T9A4.
ACCESSION AF096375
VERSION    AF096375.1
KEYWORDS   GI:3695400
SOURCE     Arabidopsis thaliana.
ORGANISM   Arabidopsis thaliana.
REFERENCE 1 Arabidopsis thaliana.
AUTHORS    Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
AUTHORS    Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
TITLE      1 (bases 1 to 83922)
AUTHORS    Washington University Genome Sequencing Project.
TITLE      The A. thaliana Genome Sequencing Project
AUTHORS    Unpublished (1997)
REFERENCE 2 (bases 1 to 83922)
AUTHORS    Zidanic, M., McQuerry, Y. and Smith, A.
TITLE      The sequence of A. thaliana T9A4
AUTHORS    Unpublished (1998)
REFERENCE 3 (bases 1 to 83922)
AUTHORS    Waterston, R.
TITLE      Direct Submission
AUTHORS    Submitted (01-OCT-1998) Department of Genetics, Washington
AUTHORS    University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     Submitted by:
            Genome Sequencing Center
            Department of Genetics, Washington University,
            St. Louis, MO 63108, USA
            e-mail: rwlison@watson.wustl.edu

```


WENGVCPIGTVPITPVTKDALRMKSPDSNSNPOSMSKTVYPASSIDEHFAV
RTTKGRSYNGASNNINFTPSVPMQPSASRNHFOIGNEFIYGMIDKINMMLY
GTSMEVGFMSRPFKESGSEWMEGEVSPSPMPNPNMNSHYPRKGSFVSVYRLI
TTVDENNTDKVNTKRYSDSCYKVRDATETWMSVGH111G3PGCK"

gene .44127)
/gene="T9A4.10"
complement(join(41254, .41451, 42287, .42442, 42949, .43137,
43675, .43767, 43984, .44127))
/gene="T9A4.10"
/codon_start=1
/evidence=not experimental
/protein_id="F24G24.10"
/db_xref="GI:3695410"
/translation="MCMILFCYILCCSLMSHCHGVGEAKAKLSNDELEQKLELI

Query Match 35.8%; Score 282.8; DB 8; Length 81922;
Best Local Similarity 99.3%; Pred. No. 1,8e-64;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCATCAACAAAACATTCTCAATACACAAAACAAAACAAAGATTATTTCTC 60
DB 26643 ATCATCAACAAAACATTCTCAATACACAAAACAAAACAAAGATTATTTCTC 26584
QY 61 TGAAGAAAGATAGTTCTACAGCAAGATGACAGTGGACATCGGAGCCGTA 120
DB 26583 TGAAGAAAGATAGTTCTACAGCAAGATGACAGTGGACATCGGAGCCGTA 26524
QY 121 GAGGCAATTAAGACCACTAGCTTTTGTGCGTGAACCTACATCTCCGTTAAT 180
DB 26522 GAGGCAATTAAGACCACTAGCTTTTGTGCGTGAACCTACATCTCCGTTAAT 26464
QY 181 CAACATCTCCGAAACAGCTTGATCTGTTTCTCAAGGAAAAGCTTCTCTCTCT 240
DB 26463 CAACATCTCCGAAACAGCTTGATCTGTTTCTCAAGGAAAAGCTTCTCTCTCTCT 26404
QY 241 GTCTCGCAGCCGTTACCTCTCTGCTGAGCGAGCAAGACGAAGA 286
DB 26403 GTCTCGCAGCCGTTACCTCTCTGCTGAGCGAGCAAGACGAAGA 26358

RESULT 6
ATF24G24 99856 bp DNA linear JUN 27-AUG-1993
LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone F24G24 (ESSA
DEFINITION
ACCESSION AL049488
VERSION AL049488.1 GI:4538949
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 99856)
Beyan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W.,
Mayer, K.F.X. and Schellier, C.
Unpublished
2 (bases 1 to 99856)
EU Arabidopsis sequencing project.
Direct Submission
Submitted (23-MAR-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
schellier@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosome 4 can be
viewed at: <http://webseq.mips.biochem.mpg.de/proj/thal/>.

COMMENT
FEATURES
source
Location/Qualifiers
1..99856
/organism="Arabidopsis thaliana"
/variety="Columbia"

/db_xref="taxon.3702"
/chromosome="4"
1..4596
/note="Overlap to BAC F28M11, please refer to this entry
for analysis and annotation"

gene .5818
CDS join(2780, .297, 3173, .3265, 3758, .3991, 4498, .4657,
4905, .5027, 5485, .5818)
/gene="F24G24.10"
/note="Similarity to IB1C3-1 protein, Arabidopsis
thaliana, A001845"

/codon_start=1
/product="putative protein"
/protein_id="CAB39774.1"
/db_xref="GI:4538950"
/translation="MKMKIGSLPVWCMILFCYILCCSLMSHCHGVGEAKAKLSNE
DLEIQRKLELNKHTVKTCTNGERYGCVDFYKQPGLDHSLMKVHTFHRRLMSYP
EGSRIRKQTHIRKTFGRFWKNGVGRPIGTVPILVSKALLMKSPDSNSNPOSMS
KIVKPTSSNGHFFAVRTTKGPRRYNGVANNINSFNPVGPMFESFGRNHFQIGIE
FVGVGWTVRGHCYVPLCPVSGIILVSHETVPLGLTKHNDPELSI IKDKIYGHM.L
KGRSSSTKKEIGFETFRFKESFGTGVMEGEVSPASTSPMGNSHFPGSPK:DS
VYRLITVDENYGLDVMVKTFRFNSCYKVRDASRGQM"
2780..2971
/gene="F24G24.10"
/number=1
2972..3172
/gene="F24G24.10"
/number=1
3173..3265
/gene="F24G24.10"
/number=2
3266..3757
/gene="F24G24.10"
/number=2
3758..3991
/gene="F24G24.10"
/number=3
3992..4497
/gene="F24G24.10"
/number=3
4498..4657
/gene="F24G24.10"
/number=4
4658..4908
/gene="F24G24.10"
/number=4
4909..5027
/gene="F24G24.10"
/number=5
5028..5488
/gene="F24G24.10"
/number=5
5489..5818
/gene="F24G24.10"
/number=6
5819..10641
/gene="F24G24.20"
join(6885, .8915, 9402, .9554, 10285, .10641)
/gene="F24G24.20"
/note="Similarity to IB1C3-1 protein, Arabidopsis
thaliana, A001845"

/codon_start=1
/product="putative protein"
/protein_id="CAB39775.1"
/db_xref="GI:4538951"
/translation="MKMSYBSGSKMKRKHRTHTFTFWENGVCPIGTVPITPVTKD
ALRMKSPDSNSNPOSMSKTYKFASSIDDDHFAVRTTKGTSYNASNNINTFTP
SVGPMPFASARNHFOIGNEFIQVGMIDKINGMMLMGTSMEEVGFMSRPFKSSST
WMEVGEVSPSPNPNPMGNSHYPRKGSFVSVYRLITVDENNTDKVNTKRYSD
SCYKVRDATETWMSVGH111YGPCK"
8685..8915
/gene="F24G24.20"

Biochemie, Am Klopferspitze 18a, D-82152 Martinsried, FRG, E-mail: lemckes@ips.biochem.mpg.de, mayet@ips.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk

Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/> this fragment has an overlap with AtCHR1V28 at the 5' end and an overlap with AtCHR1V30 at the 3' end.

COMMENT

FEATURES

source 1..199861
/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
gene="AT4g10270"
7035..7307
/gene="AT4g10270"
7035..7307
/note="strong similarity to wound-induced protein, Lycopersicon esculentum, PIR2:S19773 contains EST gb:A1995575.1"
/product="probable wound-induced protein"
/protein_id="CAB78150.1"
/db_xref="GI:7267724"
/translation="MSTSKAMTVAVSIGVEAKDQLGCRNVIIRSYNQHRLNV KSVSQKRPSSSSVAAVTSSGSESEKAKAEESLRVYLSGWP"
7035..7307
/gene="AT4g10270"
/number=1
7450..8226
/gene="AT4g10280"
/complement(join(7450..7666,8021..8226))
/gene="AT4g10280"
/complement(join(7450..7666,8021..8226))
/gene="AT4g10280"
/note="similarity to predicted protein, Arabidopsis thaliana"
/codon_start=1
/product="putative protein"
/protein_id="CAB78151.1"
/db_xref="GI:7267725"
/translation="WGNHFLTSLLVAVCVCSFITKLNPKKALIVSSDSEIFTE IHGKILROASDTKLALQGVASWPKMGAPSKPFMEFKTETIYFNEGKVKVAVGID EEEETFEIGKGVVFPKDMKVWEITEAVKKQYSLSE"
exon complement(7450..7666)
/gene="AT4g10280"
/number=1
intron complement(7667..8020)
/gene="AT4g10280"
/number=1
exon complement(8021..8226)
/gene="AT4g10280"
/number=2
cRNA 10468..10619
/product="cRNA-Ser"
/note="codon recognized: GGA"
12188..12418
/gene="AT4g10290"
complement(12188..12418)
/gene="AT4g10290"
complement(12188..12418)
/gene="AT4g10290"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB78152.1"
/db_xref="GI:7267726"
/translation="MPKRFPMKFKTTMTFVEGKLVKVEDHKEGAELEVAQDLV VEPQMVNFVDIEVKKRYRESEELP"
exon complement(12188..12418)
/gene="AT4g10290"

gene 5794..18176
/gene="AT4g10300"
/in(16794..16853,17972..18176)
/gene="AT4g10300"
/note="similarity to predicted protein, Arabidopsis thaliana contains EST gb:A1823169.1"
/codon_start=1
/product="putative protein"
/protein_id="C2578253.1"
/db_xref="GI:7267727"
/translation="KQICGVNVTSTINPYLTRSKPKYNSRRPSMAALVPAEST-KL IITEKNIPESKLTQLGVASWPKMGCPSPKPTYSAKETCYLLQGVKTVFQNSDGG EIEIADGPFVFPKMGSCITWDSVAVADKHQPE"
5794..16993
/gene="AT4g10300"
/number=1
15994..17971
/gene="AT4g10300"
/number=1
17972..18176
/gene="AT4g10300"
/number=2
14238..27814
/gene="AT4g10310"
/in(24238..25310,26586..26819,27652..27814)
/gene="AT4g10310"
/note="similarity to potassium transport protein TRK1, Saccharomyces cerevisiae, PIR2:J00466"
/codon_start=1
/product="potassium uptake transporter-like protein"
/protein_id="CAB78154.1"
/db_xref="GI:7267728"
/translation="MDRVVAKIARIQSRLTKLSHFLLFYFLFSSFLGLAKLTK RTTSRPHDFDLFTSVAITVSSMSVTDMVEFSNLTFLTLPFLOGEIPLSFNL VYSYTFKFPFNKIRHILIGSYNSDSIEDRCDETVDVDEGLIKIDEASACTVAV LSVHLVTNLVGSVLLVYVNFVKARAVLSKESPLTFSPVTTSTFPAICGVPTN ENMIFRNKSLMLIPVYLVKNTLPFCFLVLLMGVYKTRDEYGLDKNNKMG ISHLLSVRLCVLGVTVLGFLITLQLPFCAEFWSELEGMSYEKLVGSLFOVVISR HIGETVDSLTSPALVLFILMYLPPTLFPMTPEQTKIEKGDDSDSENGK/KK SGLIVQLSFLTICIFLISITERONLQDDPINFVNLITLEVIRIINKISFVMAVEWE LVSILPRTAVNGWTSLSAWAKTRVWGQDDOVQWENSY"
24238..25310
/gene="AT4g10310"
/number=1
25311..26585
/gene="AT4g10310"
/number=1
26586..26819
/gene="AT4g10310"
/number=2
26820..27651
/gene="AT4g10310"
/number=2
27652..27814
/gene="AT4g10310"
/number=3
27815..36739
/gene="AT4g10320"
complement(join(29756..30244,30328..30393,30579..30721, 30870..30937,31013..31076,31128..31311,31418..31506, 31653..31716,31808..31911,32003..32140,32227..32352, 32462..32671,32770..32853,32977..33136,33651..33784, 33863..33952,34221..34379,34479..34557,34914..35011, 35156..35259,35441..35962,36051..36116,36206..36739))
/gene="AT4g10320"
complement(join(29756..30244,30328..30393,30579..30721, 30870..30937,31013..31076,31128..31311,31418..31506, 31653..31716,31808..31911,32003..32140,32227..32352, 32462..32671,32770..32853,32977..33136,33651..33784, 33863..33952,34221..34379,34479..34557,34914..35011, 35156..35259,35441..35962,36051..36116,36206..36739))

/gene="AT4g10320"
/note="strong similarity to isoleucine--tRNA ligase, Hcra
sapiens, PIR2.159314
Contains Aminoacyl-transfer RNA synthetases class-1
signature AA49-60; Prokaryotic membrane lipoprotein lipid
attachment site AA78-798; Prokaryotic membrane lipoprotein
lipid attachment site AA125-1255"
/codon_start=1
/product="isoleucine-tRNA ligase-like protein"
/protein_id="CAI7815.1"
/db_xref="GI:7267729"
/translation="MEVEGEKEFFPPOEDVVSFTETETATKRETPETIT
YDGPATLPHGHLAGITIDIVTROTGHVTRFSGHGVPEVETDELIN
IKARDEIVMGIDKNEECRSIVIRYVMEVETVRCRMIDENDKIVPEFMSV
MWYFQLEKKNLYRQFKMYPYSGKTPLSNFGQYKEVPELWVFPVIGDQD
NAFVATITPTLPSNLALCNAAKVVYKKNIGVYVASERLSAFTDPSAK
LSNGPADITKKNPKAKPESADSEVELEKNGALVYKVEFVPSPPSSA
FRVADYVTDGSGIVHCAFGEDVRYLTKIKKENIVVAVDDQFTETI
TFESGVYDADDDITIEVAKAGLVKGTSTHSTPFCNSDTLITRAYPSMVEE
QLEKLSNEOTENVGVKDKRFHNMENRMAISRFRKQTPVIMISDGEV
VIMDSVEKLEKSGVFPDLRHHIDHITPSSRDEFPVLRVEDVDFMSGXP
VAYIHPFENKELFEKNPFGDVAGLDOTRGMFTLVVSTALFEKPAKMLICVL
VLAEDGKMAKRLNYPPLVEIDEGADAVLYINSPVVAFPLEKKEGVYKX
DVLPMVAVRPLVONAKRLTEGGVFPVPLDATIONAILDOMHSATQSIVET
EEMDAYRLVTVPRLLKPLDNLINIVRPNRKLKRGCEDDSTALSTIPNLISC
KVAFPTPTFTLTONLRKACKSESVHYSIPRRCGEGRILESTRANKIDL
ARNIRERKPLPLKEMIVVHPADPLNDITGVLRVLEELNVSILVPCDITLY
ASLKAPDFSVLAKRKGSMGLVAKVEKMSKDLAEAEVITANHLKEIDIV
SHAMFSCIELAIRIVAFRRPDLKENEDSGDDVLIIDLADDSLVEAGARAI
VNRIOKLKSGLEPTFEVYFOSLDEDESVKQVLSGKNDISGSLTSSIM
PSHAVIADETFPKETSDESVKVPKLSYSLARPAKFNENAVALVSGSFPMH
NRTSLISRLFLNVIYISRCETICNANSLVYKGFEEIVPSRRENGSRVLL

Query Match 35.8%; Score 282.8; DB 9; Length 13936;
Best Local Similarity 39.3%; Pred. No. 2.1e-64;
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATCTACAAACAAACATTCTCATATACACAAACAAACAAAGAGTTAAATCTC 63
|||||
6966 ATCTACAAACAAACATTCTCATATACACAAACAAACAAAGAGTTAAATCTC 7025
|||||
51 TGAAGAAAGATGAGTTTACAGCAAGATGACAGTGGCAGTACGATCGGACCCGTA 120
|||||
7026 TGAAGAAAGATGAGTTTACAGCAAGATGACAGTGGCAGTACGATCGGACCCGTA 7085
|||||
121 GAGGATTAAGACCAAGTACGCTTTCGCTGAGACCTACTACTCCGCTGATAT 180
|||||
7086 GAGGATTAAGACCAAGTACGCTTTCGCTGAGACCTACTACTCCGCTGATAT 7145
|||||
181 CAACATCTCCGGAACAAGTATGTTCTTCAAGGAAAAGTTCTCTCTCTCTCT 240
|||||
7146 CAACATCTCCGGAACAAGTATGTTCTTCAAGGAAAAGTTCTCTCTCTCTCT 7205
|||||
241 GTCTCCGACGCTTACCTCTCTCTGATGAGAGTAAAGACCAAGA 286
|||||
7206 GTCTCCGACGCTTACCTCTCTCTGATGAGAGTAAAGACCAAGA 7251
|||||

RESULT 9
AP005115 139336 bp DNA linear HTG 20-APR-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone P0700F06,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AP005115
VERSION AP005115.1 GI:20219003
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group) (cultivar::japonbare) DNA,
clone:P0700F06.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euharoidae; Oryzaceae; Oryza.
REFERENCE 1 Sasaki, T., Matsumoto, T. and Yamamoto, K.

AUTHORS Sasaki, T., Matsumoto, T. and Karayose, Y.
TITLE Oryza sativa japonbare (JAS) genomic DNA, chromosome 2, PAC
clone:P0700F06
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 139336)
AUTHORS Sasaki, T., Matsumoto, T. and Karayose, Y.
TITLE Direct Submission
JOURNAL Submitted (18-APR-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 139336
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="japonbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0700F06"

BASE COUNT 40464 a 28725 c 29241 g 40806 t 100 others

Query Match 15.9%; Score 125.2; DB 2; Length 139336;
Best Local Similarity 63.3%; Pred. No. 2.8e-22;
Matches 190; Conservative 1; Mismatches 109; Indels 0; Gaps 0;

1 324 AAGGAAAGATGAGTTTACAGCAAGATGACAGTGGCAGTACGATCGGACCCGTA 383
|||||
DB 62759 AAGGATGAGTTTACAGCAAGATGACAGTGGCAGTACGATCGGACCCGTA 62700
|||||
CY 384 AAGGATGAGTTTACAGCAAGATGACAGTGGCAGTACGATCGGACCCGTA 443
|||||
DB 62699 AAGGATGAGTTTACAGCAAGATGACAGTGGCAGTACGATCGGACCCGTA 62640
|||||
CY 444 CTACGCTCCGGAACAAGTATGTTCTTCAAGGAAAAGTTCTCTCTCTCTCT 503
|||||
DB 62639 CTACGCTCCGGAACAAGTATGTTCTTCAAGGAAAAGTTCTCTCTCTCTCT 62580
|||||
CY 504 TAGATGATGAGACCAAGATGTTCTCTCTTGAAGGTTCTACTGAGAGAAGA 563
|||||
DB 62579 TAGATGATGAGACCAAGATGTTCTCTCTTGAAGGTTCTACTGAGAGAAGA 62520
|||||
CY 564 GATCATACTCTTATGATTTGGAGACCAATTGAAGTAAAGTCTCTGTTGGCCG 623
|||||
DB 62519 GATCATACTCTTATGATTTGGAGACCAATTGAAGTAAAGTCTCTGTTGGCCG 62460
|||||

RESULT 9
AP004054 133158 bp DNA linear HTG 21-MAR-2002
LOCUS AP004054/C
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone
OJ1249.F12, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AP004054
VERSION AP004054.1 GI:15208422
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group) (cultivar::japonbare) DNA,
clone:OJ1249.F12.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euharoidae; Oryzaceae; Oryza.
REFERENCE 1 Sasaki, T., Matsumoto, T. and Yamamoto, K.

TITLE
 Oryza sativa nipponbare(GN3) genomic DNA, chromosome 2, BAC
 clone:OJ1249 F12
 JOURNAL
 Published Only in Database (2001)
 REFERENCE
 2 (bases 1 to 131158)
 AUTHORS
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE
 Direct Submission
 Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
 Agrobiological Resources, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasakignias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 COMMENT
 The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source
 1. 131158
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1249 F12"
 BASE COUNT 38255 a 27854 c 27875 g 39121 t 53 others
 MGIN
 Query Match 15.5%; Score 122; DB 2; Length 131158;
 Best Local Similarity 63.3%; Pred. No. 2e-21;
 Matches 190; Conservative 1; Mismatches 109; Indels 0; Gaps 0;
 324 AAGGAAAAAGATGATTCACCCGACGACCTAGCCATTACAGGCAACCCAGCATC 383
 19147 AAGAGATGGCGGGAGCTGACGACGGCGGCGACCTCGGCGCTACGACGGAGCCATC 19158
 384 AAGCGGATCTACTGCGGATCAAGCGCGTGTGATGATCACCACCGGAATCCT 413
 19087 GAAGCGATCTTCTCTCCGCGGAGAGCTTACGAGTACCTCGGAGCCGCTT 19088
 444 CTACGGCTCCGAGGCGATTACTGATGTCGCGGAAAAAGACCGAGAGACTTTGG 503
 19027 CTACGGCGCGCGCGGCGCTACCGCGCTTGGCGCGGAGAGCGCGCGCTCG 18968
 504 TAGAGTACTAAGACGAGAGAGATGTCTCTCTTCTTGAGAGTCTACCTAGGAA 563
 18967 CAGAGATGTCAGAGGAGAGCGGCGAGTCTCCGCGACCTCCGCGCTCTCCAC 18968
 564 GATCAATCTCTTAATGATTGGAGACCAATTGAGCTAAGATCTGTTTGGCGG 623
 18907 GCTCGCGCTCTCCGCGGAGAGAGAGTTCAGGCGCAAGTACCCCTCTCCCGCG 18848
 RESULT 10
 1994 83922 bp DNA 1linear PLN 06-OCT-1998
 DEFINITION Arabidopsis thaliana BAC 1994.
 ACCESSION AF066373
 VERSION AF066373.1 GI:3695400
 KEYWORDS
 ORGANISM
 Arabidopsis thaliana.
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 83922)
 REFERENCE
 Washington University Genome Sequencing Center.
 The A. thaliana Genome Sequencing Project

JOURNAL
 Published (1997)
 REFERENCE
 1 (bases 1 to 83922)
 AUTHORS
 Lander,E., Linton,M., McQuerry,Y. and Smith,A.
 TITLE
 The sequence of A. thaliana 1994
 JOURNAL
 Published (1998)
 REFERENCE
 3 (bases 1 to 83922)
 AUTHORS
 Waterston,R.
 TITLE
 Direct Submission
 Submitted (01-OCT-1998) Department of Genetics, Washington
 University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT
 Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63108, USA
 e-mail: rwlson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by
 M. Lohr, Cold Spring Harbor Laboratories, and fingerprinted
 by M. Marra, Mashu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we provide a small overlap between
 neighboring submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded or sequenced with an alternate
 chemistry; an attempt was made to resolve all sequencing problems,
 such as compressions and repeats; all regions were covered by
 sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 3' clone is T17F16. Actual start of this clone is at base
 position 1 of 1994; actual end is at 83922 of 1994.

NOTES:

Coding sequences below are predicted from computer analysis, using
 the program GeneFinder (P. Green and L. Hillier, ms in preparation).

FEATURES
 source
 1. 83922
 /location/Qualifiers
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="IV"
 /map="unknown"
 /clone="1994"
 436 3161
 gene="1994.1"
 3161 486 665 756 839 938 1147 1257 1332 1469 1606
 1568 1801 1893 1944 2103 2191 2293 2481 2533 2596
 2672 2739 2908 3030 3126 3161
 gene="1994.4"
 note="similar to [isoleucyl-tRNA synthetases]"
 codon_start=1
 evidence=not experimental
 protein_id="A062806.1"
 /db_xref="GI:3695406"

```

CDS
    /gene="T9A4.5"
    /complement(join(5712..5939,6786..6918,8299..8989,
    9346..9371))
    /gene="T9A4.5"
    /note="similar to potassium transport proteins"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62807.1"
    /db_xref="GI:3695407"
    /translation="MDRVVAKIKIRSQLTKLSLFLFYFIYFLFFSLGFLAKIKR
    PRITSRPHDLEFFTSVAITVSSMTVMVEVFNTOILFILIMPLGGELFTSLK
    YSFTFNSIEDRCDEVTVDYREGLIKIDEPASCKLYSVLSYHLVTLGSSVLLIY
    VMFVKTAPDVLSSKEISPLTFSTVTFVTFACGCVPRNKKIIFRRSGLIMLIIQ
    YLKKITLPGCVLTIKGNKIKRDEYVILKNNKSHLSLRCLTLLGVTTLG
    FLITOLFFCAFETWSESDEGSSSEKLVGSLPQVNSRHGTETIVDLSTLSPALVQ
    FLIMIGLIVQSLFTLTCIFLISITERQNLQSGPFWNLITLEVRVPCMGSAVNS
    VGFITGYSCERRVDISDGGCKDASVGFAGRSPMGKFVLIIVMEYGRFKQTKASGA
    WLPSSSS"
    /complement(26302..26574)
    /gene="T9A4.6"
    /complement(26302..26574)
    /gene="T9A4.6"
    /note="contains similarity to Solanum lycopersicum
    (tomato) wound-induced protein (GB:X59882)"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62808.1"
    /db_xref="GI:3695408"
    /translation="MSSTKANTVAVSTAVNALDQGLCRKNYIIIPSYNOHLENNV
    RSVQGRFSSSSVSAVTSGESEKAKKAEESLRTVWLSQWGNP"
    28609..29893
    /gene="T9A4.3"
    /join(28609..28785,28860..29270,29345..29479,29558..29674,
    29759..29893)
    /gene="T9A4.3"
    /note="contains similarity to the pKb family of
    carboxylate kinases (Pfam: PF00294, E=1.6e-75)"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62803.1"
    /db_xref="GI:3695403"
    /translation="MANPLIVSFGEMLIDFPVDTISGSLSESTGFLKA;VAPAVNA
    CATIKLGKSAPIGKFGDDEFGHMLVNLILKRVNSSEVCTDINARTALA;VILKKG
    EREFMFRNPDSADMLKESLEIKAKIFHYGISLISEPCRTAHMAAMTADDA
    GVLASDPVNRPLPMPSTEAIEGKISINEMDILKVSDEVTFLIGDMEKDVVIS
    LMHDKLKLIVTDGKCRVYTKFKGRVGVAVAVDTTACGDSFYGALVSLGKDG
    SIIDDEGKLKEALAFNACGAVCTTQKGAIPALPTPADAGLTKSKSK"
    /complement(30255..30842)
    /gene="T9A4.7"
    /complement(30255..30842)
    /gene="T9A4.7"
    /note="contains similarity to heat shock hsp20 proteins
    (Pfam: PF00011, E=1.2e-46)"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62802.1"
    /db_xref="GI:3695402"
    /translation="MKHLSTLFIGALLLNKITSSEGLSSALETTPGSLSDMLD
    REFDPKILERIPLGERDTSVALSPAVDMKETAEGHEIILDIPGKDEVEKLEVE
    NGYLVSGERKEEEKGDQMRVERSYGKFWPKLPDPTDWSVAKLENGVLTIN
    LTYLSPKXGPRPVVIAAEEDOTAKISSSEKEL"
    31853..32412
    /gene="T9A4.2"
    /join(31853..32053,32125..32412)
    /gene="T9A4.2"
    /note="contains similarity to Arabidopsis thaliana
    salt-tolerance protein (GB:X95572) and CONSPANS-like 1
    proteins"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62805.1"
    /db_xref="GI:3695405"
    /translation="MKIOCEVCTERAEAEVLCCSDEAVLCCGECIDIKVFAVLLFGPRHR

```

```

VALQKDAASATIASGAPLDCIDCOERKGFEECLDEBAMLCHNDENAIHTCHNSHOFPLS
GYVSDSDSLTENSECSTSPSSSEYTOIQSVKSVLSNOYSKEETEDAGSGEIVHNPVIL
SP"
    /complement(34194..35118)
    /gene="T9A4.8"
    /complement(join(34194..34922,35026..35118))
    /gene="T9A4.8"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62809.1"
    /db_xref="GI:3695409"
    /translation="MTSLIYAVSKLRGVNOIENKNSPGASEEDILNOKALLTOYER
    YKRGFKPDHWPPLIKGIEKRPANDNMKTTPAPFGEDVDTSSSFSINTESSPSGMS
    IDLNMCEDEMFSLSESPMGILKXAKQSEEPQKOLLEONDKLIKATIKTSRNEI
    QROKIEVARKEENKILFADLNLISDPSRAVNERKRIIEKAQTNQHEDEGGO
    YHGSQYASVSHYSLFPHGQVOGEPPQGDKNSSPNNODPFOYNYLSTGNP"
    /complement(36300..39184)
    /gene="T9A4.9"
    /complement(join(36300..36656,37387..37539,38026..38214,
    38746..38838,38999..39184))
    /gene="T9A4.9"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62813.1"
    /db_xref="GI:3695413"
    /translation="MKLGCLPMTYCNILACVIIGSLIMSHQGVIEAKTLKSNEDL
    EIEHKLILNKPAPKIVKTINGERQCVDFYQPLDSSMKNTFHHKTRNTTFGH
    FWENGVCPIQVPIPRVTKDLNLRKMSDSNPSQSSWSEKTYKPAISSIDHHFAVY
    FTTKGRISYNGAEMINIFTPEVSGPWFASMSHFOINIEIOGMIDIKINGNMWLM
    GTSWEEVGFPPSSRFKESGIVIEWGVEVTSPPNPMGNSSHYPKSGPKYDSVRLI
    TTVDENYNDIKYNTERYSDSCYKRDATETFMHSHVHLIIYGGPGCK"
    /complement(41254..44127)
    /gene="T9A4.10"
    /complement(join(41254..41451,42287..42442,42949..43137,
    43675..43767,43984..44127))
    /gene="T9A4.10"
    /codon_start=1
    /evidence=not experimental
    /protein_id="AAC62810.1"
    /db_xref="GI:3695410"
    /translation="MGLILFCYILCCSLIMSHQGVEMAAVAKLSNEDLEIQLELI
    Query Match 14.44: Score 113.8; DB 8; Length 83922;
    Best Local Similarity 70.44; Pred. No. 2,9e-19;
    Matches 171; Conservative 0; Mismatches 62; Indels 10; Gaps 1;
    QY 1 ATCATCAACAAACAAATTCATACACAAACAAACAAACAAAGAGTTAACTGTC 60
    Db 27843 ATCATCAACAAACAAACAAATGAAACCAATCACTTAAACAAAGAA-----A 27892
    QY 61 TGAAGAAAGATGAGTTCTACAAAGCAAGTCGACGTGAGCATCGGAGCCGTA 120
    Db 27893 CAAGAGAAATGAGCTCTCGACAGCAAGAGCTGATGTCGCAAGCATGGACCCCT 27952
    QY 121 GAGGATTAAAGACCACTAGTCTTTGTGGTGAAGTACATCTCGGTGGTTAAT 180
    Db 27953 GAGGATTAAAGACCAACTAGGCGTGTGCTTGGACATAGCATTCGATCCGAAAT 28012
    QY 181 CAACATCTCCGAAACAGTATGATGCTTTCAGAGGAGAAAGTCTCTTCGCTTCT 240
    Db 28013 CAGTATCTACCCAAACATTAAGTTCGTCGCAAGCTTAAGAGCTCTTCTCCTACA 28072
    QY 241 GTC 243
    Db 28073 ATC 28075
    RESULT 11
    AT24624/c 99856 bp DNA linear PLN 27-AUG-1999
    LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone F24G24 (SSA
    DEFINITION Project).
    ACCESSION AF049489

```

```
VERSION      AL049468.1  GI:4538949
KEYWORDS
SOURCE       Arabidopsis thaliana.
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicot;
              Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.s.
REFERENCE    1 (bases 1 to 99856)
AUTHORS      Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Wekes, H.W.,
              Mayer, K.F.X. and Schellier, C.
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 99856)
AUTHORS      EU Arabidopsis sequencing project.
JOURNAL      Direct Submission
              Submitted (23-MAR-1999) MIPS, at the Max-Planck-Institut fuer
              Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
              schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
              Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
              Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
              E-mail: michael.bevan@bsrc.ac.uk
              Information on performance of analysis and a more detailed
              annotation of this entry and other sequences of chromosome 4 can be
              viewed at: http://www.mips.biochem.mpg.de/proj/thai/.
FEATURES
  source
    1..99856
       /organism="Arabidopsis thaliana"
       /variety="Columbia"
       /db_xref="taxon:3702"
       /chromosome="4"
  misc_feature
    1..4596
       /note="overlap to BAC F28M11; please refer to this entry
       for analysis and annotation"
  gene
    2780..5818
       /gene="F24G24.10"
       join(2780..2971,1173..3265,3758..3991,4498..4557,
       4909..5027,5489..5818)
       /gene="F24G24.10"
       /note="similarity to IB1C3-1 protein, Arabidopsis
       thaliana, A011845"
       /codon_start=1
       /product="putative protein"
       /protein_id="CAB39774.1"
       /db_xref="GI:4538950"
       /translation="MKMKLGSLPYMCNILFCVILCCSLMSHCQVGEAKALKSNE
       DLEIQLKELINKHTVKIILKCTNGERYGCVDPYQPLDHSILKNHIFHHKRLMSYP
       ESKIKQTHINKTFGHFKMGVGPIDGTVPILVSKKALLKMSFEGD;SNPSSMS
       KTYKPTSSNGGHHFAVRRTKGKPRRVNGVMNINSPNPGVMEFSAGRWHEQIQNE
       FVOGMYVRGHCYNPLCPVGGIILVSHVTPGLLRKNDPELSIIKDKIKYQHWLL
       MNSSSTMKKEIGFWPTHRPFSQTCVENGCEVYSPASTSPMGNSHFFPKSPKIDS
       VRLITTTDENGLDMVYKNTFRSNSCYKXKQSRFGOM"
  exon
    2780..2971
       /gene="F24G24.10"
       /number=1
  intron
    2972..3172
       /gene="F24G24.10"
       /number=1
  exon
    3173..3265
       /gene="F24G24.10"
       /number=2
  intron
    3266..3757
       /gene="F24G24.10"
       /number=2
  exon
    3758..3991
       /gene="F24G24.10"
       /number=3
  intron
    3992..4497
       /gene="F24G24.10"
       /number=3
  exon
    4498..4657
       /gene="F24G24.10"
       /number=4
  intron
    4658..4908
       /gene="F24G24.10"
       /number=4
  intron
    4909..5027
       /gene="F24G24.10"
       /number=5
  intron
    5028..5489
       /gene="F24G24.10"
       /number=5
  exon
    5489..5818
       /gene="F24G24.10"
       /number=6
  gene
    5819..10641
       /gene="F24G24.40"
       join(5819..5915,9402..9554,10285..10641)
       /gene="F24G24.20"
       /note="similarity to IB1C3-1 protein, Arabidopsis
       thaliana, A011845"
       /codon_start=1
       /product="putative protein"
       /protein_id="CAB3975.1"
       /db_xref="GI:4538951"
       /translation="MKMSYSEGSKMKRKTNRNTFGHFWEWGCGPIGTVP;ITKD
       ALLRMSFSDNSNPQSSWSKTYKPAASIDHHFVAVTTKTRSYNKAASKNIN;FTP
       SVGMPQFSAERHAFQIGNEFIQVGMIDKIMWMLIMGMEVGVFWSSEKSGT
       MVEGCEVYSPSPNPMPGSHYKPSKPSVDSYVRLITVDENYITDVTAKTEYSD
       SCYKVRDARETFMSVGHLLIYGPGCK"
       8685..8915
       /gene="F24G24.20"
       /number=1
       8916..9401
       /gene="F24G24.20"
       /number=1
       9402..9554
       /gene="F24G24.20"
       /number=2
       9555..10284
       /gene="F24G24.20"
       /number=2
       10285..10641
       /gene="F24G24.20"
       /number=3
       10642..11823
       /gene="F24G24.30"
       join(10642..11823,11915,12019..12747)
       /gene="F24G24.30"
       /codon_start=1
       /product="hypothetical protein"
       /protein_id="CAB39776.1"
       /db_xref="GI:4538952"
       /translation="MTSILSAVSKLRCGVNOIENKPSGASEEDILNQAKMLLTQYER
       KRGKPFDFVMPILKGIKEKANDNMKTPPAFGEGRDVTSSSPSINTESSPFGNS
       TLNDSQDAFSLSSRPYGLKKARKQOSEQFQLCEKDKLIKAITKGTCT;NRI
       CQKIKIEMARKKEENKILFADLNSISDPSRAVENERKILKQALNCHDEDE;SQ
       YGSGYRASHYQESLFGKQVGEFQDQEDKRSPPNQSDFTQYNTILSGGNPF"
       11823..11915
       /gene="F24G24.30"
       /number=1
       11916..12018
       /gene="F24G24.30"
       /number=1
       12019..12747
       /gene="F24G24.30"
       /number=2
       12748..15088
       /gene="F24G24.40"
       complement(join(14529..14816,14888..15088))
       /gene="F24G24.40"
       complement(join(14529..14816,14888..15088))
       /gene="F24G24.40"
       /note="strong similarity to zinc-finger protein R2931,
       Cry2a sativa, PIR3:J0116"
       /codon_start=1
       /product="zinc-finger-like protein"
       /protein_id="CAB39777.1"
```

```

/db_xref="GI:4538953"
/translation="MKIOCEVCEKAFAVLCSDPAVLCKPCDIKXVHEANXLFQPHR
VALQKDAASATTAGPACDIICERKGVFCLEDRAPALCMCEEAHTICRQRELLS
GVQVSDQSLTENSECSFSFSEITVQISXLSLSQVSSEELANGNGELVHRCPSVILL
SP"
complement(14529..14816)
/gene="F24G24.40"
/number=1
complement(14817..14887)
/gene="F24G24.40"
/number=1
complement(14888..15088)
/gene="F24G24.40"
/number=2
/number=2
16099..16686
/gene="F24G24.50"
16099..16686
/gene="F24G24.50"
/feature="contains EST gb:AA597970"
/codon_start=1
/product="heat shock protein 22.0"
/protein_id="CAB39778.1"
/db_xref="GI:4538954"
/translation="MMKLLSIFFGALLGNIKTSGSLASLETTPTPSLSDMLD
RFPDEFILERIPLGLERDITVALSPARVFMKETAQHEIMIDIPLKXDEVKVEE
NGVLEVGSRKREBEKKGDWRVERSYGKFMQFLPDVWVSWKALEVCVLTIN
LTKLSEKVKGRPVNIAAEEDOTAKISSESEKEL"
16099..16686
/gene="F24G24.50"
/number=1
17048..18332
/gene="F24G24.60"
complement(join(17048..17182,17267..17333,17462..17596,
17671..18081,18156..18332))
/gene="F24G24.60"
complement(join(17048..17182,17267..17333,17462..17596,
17671..18081,18156..18332))
/gene="F24G24.60"
/feature="strong similarity to fructokinase, lycopersicon
esculentum gb:U62329
contains pfkb family of carbohydrate kinases signatures,
pfkb_kinases_1 [GGAPNVAICATIKLGGSAFIFKFG], pfkb_kinases_2
[DTGAGDSFVGAFI]"
/codon_start=1
/product="fructokinase-like protein"
/protein_id="CAB39779.1"
/db_xref="GI:4538955"
/translation="MAHTPLVSGFGEHLIDFVDTGVSLSAESTGFLKAPAGAPLVVA
CAITRLGGKSAFIFKFGDDEGHMLVNLKRGVNSEGVCEDINARTALAFVTLKXDG
14.4%; Score 113.8; DB 8; Length 5986;
Bes Local Similarity 70.4%; Pred. No. 3e-19;
Mat 95 173; Conservative 0; Mismatches 62; Indels 10; Cys 1;

```

```

DB 18666 ATC 15866
RESULT 12
LOCUS ATCHRIV29/c 199861 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 29.
ACCESSION AB161517
VERSION AB161517.2 GI:7267723
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 155039)
Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and
Mayer,K.F.X.
JOURNAL Unpublished
2 (bases 147497 to 199861)
Medler,H., Medler,E., Wambutt,R., Mewes,H.W., Lemcke,K. and
Mayer,K.F.X.
JOURNAL Unpublished
3 (bases 1 to 199861)
EU Arabidopsis sequencing project.
Direct Submission
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemckemips.biochem.mpg.de,mayermips.biochem.mpg.de/Proj/thal/
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/Proj/thal/
this fragment has an overlap with ATCHRIV28 at the 5' end and an
overlap with ATCHRIV30 at the 3' end.
Location/Qualifiers
1..199861
/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
7035..7307
/gene="AT4G10270"
7035..7307
/gene="AT4G10270"
/feature="strong similarity to wound-induced protein,
lycopersicon esculentum, PIR2:SI9773
contains EST gb:A1995575.1"
/codon_start=1
/product="probable wound-induced protein"
/protein_id="CAB78150.1"
/db_xref="GI:7267724"
/translation="MSSTSKAWTVAVSAGVAVLKQGLGRNNYILRSVNOHLRNV
RSVSQGRFSSSVSAVAVTSGESEKAKKAEESLRVWYLSGCGPR"
7035..7307
/gene="AT4G10270"
/number=1
7450..8226
/gene="AT4G10280"
complement(join(7450..7666,8021..8226))
/gene="AT4G10280"
complement(join(7450..7666,8021..8226))
/gene="AT4G10280"
/feature="similarity to predicted protein, Arabidopsis
thaliana"
/codon_start=1
/product="putative protein"
/protein_id="CAB78151.1"
/db_xref="GI:7267725"
/translation="MGNHFLTSLDLTVVCVSPITTKLPKPAIVSVSDSEIPTE
IIGVILIQASDTKLALQGLVAVSWPKMEGASFPFMEKRTETITTFEGKAYNVDSID

```

```

exon      EEENFE1GKGVVVFPRKDKVWEITEAVKQVSI EF"
           complement(7450..7666)
           /gene="ATg10280"
           /number=1
intron    complement(7667..8020)
           /gene="ATg10280"
           /number=1
exon      complement(8021..8226)
           /gene="ATg10280"
           /number=2
tRNA      10468..10619
           /product="tRNA-Scr"
           /note="codon recognized: GGA"
           /number=1
gene      12188..12418
           /gene="ATg10290"
           complement(12188..12418)
           /gene="ATg10290"
           complement(12188..12418)
           /gene="ATg10290"
           /codon_start=1
           /product="hypothetical protein"
           /protein_id="CA878152.1"
           /db_xref="GI:7267728"
           /translation="MPKRFPMKRTETWVFEQKLVKVEDHMKGEALEFFVAGDLY
           VFPDMVFEVDIEVKRYRESEIESELP"
           complement(12188..12418)
           /gene="ATg10290"
           /number=1
gene      16794..18176
           /gene="ATg10300"
           /join(16794..16993,17972..18176)
           /gene="ATg10300"
           /note="similarity to predicted protein, Arabidopsis
           thaliana
           contains EST gb:A1823169.1"
           /codon_start=1
           /product="putative protein"
           /protein_id="CA878153.1"
           /db_xref="GI:7267727"
           /translation="MGICGVNVTISINPLYTRSNKPYRRPSSMLAIEAETEL
           GITLEKPKESKLTQLGVRSMWPKGCPSPKFPMTYSAKETCYLLOGKRVVPMNSDGC
           VEIADGFVVPKMSCTNDVSAVVKHQFE"
           16794..16993
           /gene="ATg10300"
           /number=1
           16994..17971
           /gene="ATg10300"
           /number=1
           17972..18176
           /gene="ATg10300"
           /number=2
           24238..27814
           /gene="ATg10310"
           /join(24238..25310,26586..26819,27652..27814)
           /gene="ATg10310"
           /note="similarity to potassium transport protein TKR1,
           Saccharomyces cerevisiae, PIR2:U00466"
           /codon_start=1
           /product="potassium uptake transporter-like protein"
           /protein_id="CA878154.1"
           /db_xref="GI:7267728"
           /translation="MDRVVAKIAKRSQTLKRSIFFLYFIYFSLFSLAKITK
           PRTRSPHDEDFEVSATVSMTVMVEVSTQLIFLILKLOGEIEFTFLW
           VVSFTFEPHKKIRILIGSYSDSIEDCDVETVIDRGLIKIDERSKCLYSV
           VLSHLYTNVGSVLLVNVFKARDVLSKETISPLTFSEFTTTFACGFPVN
           ENMIFRKNGLIWLIPQVLMGNTLFPCLVLLMGYKITKDEYVILKHNKRG
           VSHLSVRLCVLGAVLGLIQLFPCAFMTSESGMSYEKLVGSLFQVNRG
           HTGETVIDLSTLPAIIVFLIMLYPYTLFPLEQTEKEGDDSDNGSKVKK
           SGLIVSOLFETICIFLISTERONQRPINPNVNLITLLEVINIINKISVMAVHNE
           TLVSLPQTAVGWTSMVAAKRVWGLDDDDVQWENSY"
           24238..25310
           /gene="ATg10310"
           /number=1

```

```

intron    -0311..26585
           /gene="ATg10310"
           /number=1
exon      16586..26819
           /gene="ATg10310"
           /number=2
intron    26820..27651
           /gene="ATg10310"
           /number=2
exon      27652..27814
           /gene="ATg10310"
           /number=3
gene      29736..36739
           /gene="ATg10320"
           complement(join(29736..30244,30328..30393,30579..30701,
           30870..30937,31013..31076,31128..31131,31418..31506,
           31653..31716,31808..31911,32003..32140,32227..32352,
           32462..32571,32777..32853,32977..33136,33651..33784,
           33863..33952,34241..34379,34479..34557,34914..35011,
           35156..35269,35441..35962,36051..36116,36206..36739))
           /gene="ATg10320"
           complement(join(29736..30244,30328..30393,30579..30701,
           30870..30937,31013..31076,31128..31131,31418..31506,
           31653..31716,31808..31911,32003..32140,32227..32352,
           32462..32571,32777..32853,32977..33136,33651..33784,
           33863..33952,34221..34379,34479..34557,34914..35011,
           35156..35269,35441..35962,36051..36116,36206..36739))
           /note="strong similarity to isoleucine--tRNA ligase, Hmc
           sapiens, PIR2:159314
           contains aminoacyl-tRNA transfer RNA synthetase class-1
           signature A44-50, prokaryotic membrane lipoprotein lipid
           attachment site A479-798, prokaryotic membrane lipoprotein
           lipid attachment site A41215-1225"
           /codon_start=1
           /product="isoleucine-tRNA ligase-like protein"
           /protein_id="CA878155.1"
           /db_xref="GI:7267733"
           /translation="MEVTEGKFSPPROEDVLSFWTIDAFKTLKRTENLPYI;
           WDSPFATGLPHGH1LAGITIKOTVRYQMTGHAHTRFGMDCHGLVENEIDPKL;
           HARDEVIAKIGIDVNECRSIVTRYVAEMVETIRCGMIDFNKYDTMPLPFMES;
           WAFVSQMEKLVYRGFNVPYSTGCKTLPSEFGQVKEVDEIPDEIWPVDDC;
           NEAFVAMTTTPMLPSHIALCVNAKRVYVKNKTKGYIVAEISALTPDKRKA;
           NSCPADPTKKAHPKAKGKPSAADSVYLVKKNKGYVPLPDSDDSE;
           FEVADDVYTDSDGTGYHCAFGAPGDDVRYLKKIITKGNLVAVDDGLPTER;
           FHSGRVYKADDDIEAVAKGRVLTQSTFHSIPFCRMDPTLPIVAVSWFVRE;
           FKEKLKSEQTEWVGVKDRFHNMLENDAIISRRMGTPLIMISDQEEV;
           FVYTHYPPEKKELEFEKVFJDFVAGELDQTRPMPTLWVLSALPEKAPFNLCG;
           FVADGKMAKRLNRPYFHEVIDEYGAVALVYINSVLAEBLPRKKECVGVGV;
           FVFPYFAVRFVQVNYEETCGVPPPLATIT3AAILDCGHSATSLVRF;
           EENAVRYTVVRLLKFLDNLNITVRRNRKLAGRGEGDDCHALSTLVLTES;
           ENKAPFTPTETLYQHEKACGSESESVHYCSIPRGMSEERELSVTKMILT;
           AKNIRENKLPLKTLNEMIVAPDADFNIDTGLREYVELNRLVLCNDTL;
           ASLKAEDFVGLKAKSGVAKVAKENKMDKILAFEEAGEVIAHLLKENDIV;
           SHAMPSCSEALRIIVAFKPDLEKENEIDSAGDDVVIDLRADGSLVAGAPR;
           FRIQIKRSGSGLEPTDFEVYFQSLDESEYSKQVLSQEOINIDISGTLISLV;
           FSHVLIADETPTKETSDESVKVPKYSKISLAPKILKFEKREVALYAGSTSPSS;
           NRELSLSKFLINVIYVGRKCEINRASNILIVKGSFEFEVRPSPRMKGLVK;

```

Query Match 14.48; Score 113.8; DB 8; Length 199661;
 Best local similarity 70.44; Pred. No. 3.3e-19;
 Matches 171; Conservative 0; Mismatches 62; Indels 10; Gaps 1;

```

Cy 1 ATCATCAACAAATTCATCTCATPACACAAACACAAACACAAAGAGTTAATCTC 60
    |||||
Db 5766 ATCTTCACACAAATTCATCTCATPACACAAACAAAGAA-----A 5717
    |||||
Cy 61 TGAAGAAAGATAGGTTTACACAGAAAGCATGGAACATGGCACTGACATCGAGCGCTA 120
    |||||
Db 5716 CAAGAGAAATAGGCTGCAAGCAAAACGTGATGCTGTCACAGCATCGAGCGCTT 5657
    |||||
Cy 121 GAGGATTAAGAAAGCAAGCATAGCTCTTTCGGTGAACATCACTACTCCGCTGGTTAAT 130

```

Db 5656 GAGGCACTGAAAGACCACTAGCCGTCGTCTTGAACCTAGTCGATCCGATTCGCCAAT 5597
 Qy 181 CAACATCTCCGAAACACCTTAGATCTGTTTCTCAAGGAAAGATCTCTTCGCTTCT 240
 Db 5596 CAGTATCTACGCACAACTTAGATCCGTCGTCTGACAGCTAGAGAGCTCTCTCTCATCA 5537
 Qy 241 GTC 243
 Db 5536 ATC 5534

RESULT 13
 AC130811 141382 bp DNA linear HTG 23-AUG-2002
 LOCUS Medicago truncatula clone mth2-26b4, WORKING DRAFT SEQUENCE, 9
 DEFINITION Medicago truncatula clone mth2-26b4, WORKING DRAFT SEQUENCE, 9
 AC130811
 VERSION AC130811.4 GI:22450619
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 141382)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.
 REFERENCE AUTHORS Medicago truncatula BAC Clone mth2-26b4
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 141382)
 AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 3 (bases 1 to 141382)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.
 REFERENCE AUTHORS Direct Submission
 JOURNAL Submitted (23-AUG-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 On Aug 23, 2002 this sequence version replaced gi:22297364.
 COMMENT Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKXOR

NOTE: This is a 'working draft' sequence. It currently
 consists of 9 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 2895: contig of 2895 bp in length
 2896 2995: gap of unknown length
 2996 7176: contig of 4181 bp in length
 7177 7276: gap of unknown length
 7277 13608: contig of 6333 bp in length
 13610 13709: gap of unknown length
 13710 19986: contig of 6277 bp in length
 19987 20086: gap of unknown length
 20087 36269: contig of 16183 bp in length
 36270 60069: gap of unknown length
 60070 60170: gap of unknown length
 60170 85493: contig of 25324 bp in length
 85494 85593: gap of unknown length

* 85594 112631: contig of 27038 bp in length
 * 112632 112731: gap of unknown length
 * 112732 141382: contig of 28651 bp in length.
 Location/Qualifiers
 1. 141382
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone_id="mth2-26b4"
 /clone_lib="Medicago truncatula BAC library H2"
 BASE COUNT 47478 a 23124 c 22805 g 47171 t 804 others
 ORIGIN

Query Match 11.4%; Score 90.2; DB 2; Length 141382;
 Best Local Similarity 61.1%; Pred. No. 6,4e-13;
 Matches 146; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 Qy 1 ATCATCACAACAAATCTTCATACCAAAACACAAACAAAGTTAATCTTC 60
 Db 49239 ATCATCACAATCAACAAACAACTTCAATATTTCAACAAAGGTACAACTTACA 49238
 Qy 61 TGAAGAAAGATGAGTTCTTACAAAGCAATGACAGTGCAGTGCATCGAGCCGTA 120
 Db 49239 ACAACCAACATGAGTGCAGCAACAAAGCTGGGTAGTACATCAAGCATAGAGCAGTG 49358
 Qy 121 GAGGATTAAGACCAACTAGTCTTTCGCGTGCATCTACTCCGTCGTTAAT 180
 Db 49359 GAGGATTAAGACCAACTAGTCTTTCGCGTGCATCTACTCCGTCGTTAAT 180
 Qy 181 CAACATCTCCGAAACACCTTAGATCTGTTTCTCAAGGAAAGTCTCTTCGCTTC 239
 Db 49419 CAACATCTCAAGCAATATTAGTCTCTCAAGCCAAAGAAATGCTCTCTGCTTC 49477

RESULT 14
 AF153284 847 bp mRNA linear PLN 02-JUN-1999
 LOCUS Arabidopsis thaliana putative progesterone-binding protein homolog
 DEFINITION (Atmp1) mRNA, complete cds.
 ACCESSION AF153284
 VERSION AF153284.1 GI:4960155
 KEYWORDS Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 SOURCE Arabidopsis thaliana.
 REFERENCE AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 847)
 Choi,J.H., Choi,H. and Gray,P.
 REFERENCE AUTHORS Plant homologues of mammalian putative progesterone-binding
 TITLE membrane proteins
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 847)
 AUTHORS Choi,J.H., Choi,H. and Gray,P.
 TITLE Direct Submission
 JOURNAL Submitted (19-MAY-1999) School of Biology, Georgia Institute of
 Technology, MC 0230, Atlanta, GA 30332-0230, USA
 FEATURES Location/Qualifiers
 1. 847
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="dbEST:N37902"
 /db_xref="taxon:3702"
 /chromosome="5"
 1. 847
 /gene="Atmp1"
 62...724
 /gene="Atmp1"
 /product="putative progesterone-binding protein homolog"
 /protein_id="AAD34615.1"
 /db_xref="GI:4960156"
 /translation="MALEWQTLKEAIIHAATGSPVFFALALAFAYOVISGFAS
 FDDVNRHQRASLNAQEEEPPIPOPVGVGTEETTELKQYSGDPQKLLMAIRHQIYD

gene
 CDS

© 1997 Oxford University Press 4139-4146

A fiber optic biosensor for fluorimetric detection of triple-helical DNA

André H. Uddin, Paul A. E. Piunno¹, Robert H. E. Hudson^{1,+}, Masad J. Damha* and Ulrich J. Krull^{1,*}

Department of Chemistry, Otto Maas Chemistry Building, McGill University, Montreal, Quebec H3A 2K6, Canada and ¹Chemical Sensors Group, Department of Chemistry, Erindale College, University of Toronto, Mississauga, Ontario L5L 1C6, Canada

Received May 13, 1997; Revised and Accepted August 22, 1997

ABSTRACT

A fiber optic biosensor was used for the fluorimetric detection of T/AT triple-helical DNA formation. The surfaces of two sets of fused silica optical fibers were functionalized with hexaethylene oxide linkers from which decaadenylic acid oligonucleotides were grown in the 3' to 5' and 5' to 3' direction, respectively, using a DNA synthesizer. Fluorescence studies of hybridization showed unequivocal hybridization between oligomers immobilized on the fibers and complementary oligonucleotides from the solution phase, as detected by fluorescence from intercalated ethidium bromide. The complementary oligonucleotide, dT₁₀, which was expected to Watson-Crick hybridize upon cooling the system below the duplex melting temperature (*T*_m), provided a fluorescence intensity with a negative temperature coefficient. Upon further cooling, to the point where the pyrimidine motif T*AT triple-helix formation occurred, a fluorescence intensity change with a positive temperature coefficient was observed. The reverse-Hoogsteen T·AT triplex, which is known to form with branched nucleic acids, provided a corresponding decrease in fluorescence intensity with decreasing temperature. Full analytical signal evolution was attainable in minutes.

INTRODUCTION

With recent advances in nanotechnology (1), there is an increased demand to investigate biomolecular structure and behavior (2). One particular area of interest stems from the progress in the synthesis of novel nucleic acid macromolecules. Dendrimers (3,4), circular (5) and cage oligonucleotides (6) have been synthesized and these novel compounds are finding applications in biotechnology (7,8).

- [Abstract of this Article](#)
- [Reprint \(PDF\) Version of this Article](#)
- Similar articles found in:
 - [Nucl. Acids. Res. Online](#)
 - [PubMed](#)
- [PubMed Citation](#)
- Search Medline for articles by:
 - [Uddin, A. H.](#) || [Krull, U. J.](#)
- Alert me when:
 - [new articles cite this article](#)
- [Download to Citation Manager](#)

Furthermore, there is much interest in the development of devices for rapid diagnostic assays to detect microorganisms, viruses and genetic mutations based on hybridization with immobilized nucleic acid probes. Approaches involving electrochemical (9), acoustic wave or piezoelectric (10), plasmon resonance (11,12), colorimetric sensing of non-particle aggregates (13) and fluorescence based optical fiber sensing techniques have been proposed (14-16). In these examples, identification of the analyte is based on the occurrence of Watson-Crick hybridization events, with the formation of three-stranded structures, or triplexes, being largely ignored.

Triple-helical oligonucleotides have potential use as sequence specific artificial nucleases (17), modulators of DNA-binding proteins/gene expression (18,19; for a recent review see ref. 20), materials for genomic mapping (21), and sensitive screening reagents to detect mutations within duplex DNA (22). Formation of three-stranded helices by nucleic acids is a well-known phenomenon which involves a third strand interacting with a purine rich strand in the underlying Watson-Crick DNA duplex (23,24). Two distinct classes of DNA triple-helices have been characterized which differ in the composition and orientation of the third strand relative to the Hoogsteen binding partner (25-32). Nucleic acid binding ligands can be used to identify DNA structures and morphology. For example, ethidium bromide binds to both duplexes and triplexes by intercalation (33), but there is a marked difference in the binding efficiency and fluorescence quantum efficiency between both types of complexes (34-36).

We have focused on the use of a nucleic acid binding ligand (e.g., ethidium bromide) and fluorescence transduction strategy to investigate oligonucleotide hybridization on fused silica optical fiber surfaces. Previously, we reported detection of hybridization events between fibers derivatized with single-stranded deoxyribonucleic acid and complementary DNA and RNA from solution (14). Herein, we report the use of optical biosensor technology for rapid detection of T/AT triplex formation in both parallel and antiparallel configurations. This rapid and efficient triple-helical assay may be extended to include diagnostic assays for sequence-specific duplex recognition, monitoring *in vivo* concentration of gene therapy pharmaceuticals, and for studying properties of synthetic oligonucleotides.

MATERIALS AND METHODS

Chemicals

Reagent grade solvents were purchased (BDH, Toronto, ON) and further purified or dried by standard laboratory practices. DNA synthesis reagents and decaoxyadenylate (dA₁₀) were purchased from Dalton Chemical Laboratories Inc. (Toronto, ON) and were used as received or were prepared as below. Anhydrous acetonitrile (Dalton) was predried by distillation from P₂O₅ and redistilled from calcium hydride under dry argon. Tetrahydrofuran (BDH) was predried over CaH₂, filtered and distilled immediately prior to use from sodium metal (Aldrich)/benzophenone (Aldrich). Ethidium bromide (3,8-diamino- 5-ethyl-6-phenylphenanthridinium bromide; Aldrich) was used as received. Water was double-distilled in glass, treated with diethyl pyrocarbonate (Aldrich) and autoclaved. Molecular biology grade polyacrylamide gel electrophoresis reagents and apparatus were obtained through Bio-Rad (Hercules, CA). Silica gel (Toronto Research Chemicals, Toronto, ON) had a particle size of 30-70 microns. Pre-cut fused silica optical fiber pieces with a length of 48 mm and a core diameter of 400 µm having both termini polished to within a 0.3 µm tolerance were obtained from 3M Specialty Optical Fiber (North York, Ontario, Canada) in addition to lengths of fiber having the same core material and diameter with a TECS 48 low refractive index outer cladding (0.48 numerical aperture).

Derivatization of optical fibers

Synthesis of DMT-HEG (dimethoxytritylated hexaethylene glycol). A solution of dimethoxytrityl chloride (7.1 g, 21 mmol) in dry pyridine (10 ml) was added dropwise to a stirred solution of hexaethylene glycol (HEG, 5.6 ml, 21 mmol in 5 ml pyridine) under an argon atmosphere. Stirring was continued overnight after which the reaction mixture was combined with dichloromethane (50 ml). The mixture was then shaken with 5% aqueous bicarbonate (2×90 ml) and then with water (2×90 ml) to remove unreacted HEG, pyridine and salts. The organic layer was dried under reduced pressure to yield the crude product. The product was purified by silica gel column chromatography using an eluent of 1:1 dichloromethane/diethyl ether containing 0.1% triethylamine (2.9g, 24% yield). The identity of the product was confirmed by proton NMR spectroscopy. R_f (silica gel thin-layer chromatography): 0.10 in CH_2Cl_2 /ether (1:1). ^1H NMR (200 MHz, CDCl_3) [δ]: 7.48 (t, 1H, $J = 1.8$ Hz), 7.46-7.42 (m, 2H), 7.27 (d, 1H, $J = 2.6$), 7.3 (d, 1H, $J = 3.3$ Hz), 7.1 (m, 8H), 3.79 (s, 6H), 3.64 (s, 24H). **Surface preparation of optical fibers.** The coating material was mechanically stripped from the pre-cut optical fiber pieces and the cladding dissolved by treatment with acetone. The surface of the fibers were then cleaned via treatment with 25% ammonia/30% hydrogen peroxide/water (1:1:5, v/v/v) for 5 min at 80°C followed by rinsing with 30% hydrogen peroxide. The fibers were then treated with a solution of conc. HCl/30% hydrogen peroxide/water (1:1:5, v/v/v) for 5 min at 80°C , followed by rinsing with methanol, dichloromethane and diethyl ether. **Functionalization of optical fibers with 3-glycidopropyltrimethoxysilane (GOPS).** Following a modification of the method reported by Maskos and Southern (37), optical fibers and silica gel were activated by placement into a solution of xylene/GOPS/diisopropylethylamine (100:30:1 v/v/v). The reaction was permitted to proceed with gentle agitation for 24 h under nitrogen at 80°C . The fibers and silica gel were rinsed with methanol, dichloromethane and diethyl ether. **Linkage of DMT-HEG to GOPS functionalized optical fibers.** The fibers and silica gel were then functionalized with monotritylated hexaethylene glycol (DMT-HEG) (250 mg, 0.46 mmol) in 30 ml of xylene containing a catalytic amount of sodium hydride with gentle agitation at 40°C . Silica gel samples (~10 mg) were taken from the reaction mixture daily to determine the loading of DMT-HEG, and this was presumed to indicate loading on the activated fibers. The silica gel samples were immediately washed with 10 ml portions of dichloromethane until the wash solution showed no absorption at 504 nm upon treatment with trichloroacetic acid. The GOPS-HEG-DMT functionalized silica gel samples were then dried under reduced pressure and treated with 5 ml of 5% trichloroacetic acid in dichloroethane in order to liberate the dimethoxytrityl moieties from the hexaethylene glycol chains. The absorbance (504 nm) of the resulting solution was then determined to quantitatively measure the loading of immobilized DMT-HEG. This analysis indicated that the reaction had gone to completion after 7 days. After this time, the fibers were removed from the reaction mixture, washed with dichloromethane and dried by storage *in vacuo* and over P_2O_5 overnight.

The secondary hydroxyl groups produced after reaction of the HEG linker with the epoxide moieties and all other silanols were capped via treatment with trimethylsilyl chloride in pyridine (1:10 v/v) under argon at room temperature for 16 h followed by treatment with acetic anhydride/*N*-methylimidazole/collidine in THF to prevent unwanted oligonucleotide growth at these sites (38). The fibers were then washed sequentially with pyridine, dichloromethane, methanol and diethyl ether and stored *in vacuo* and over P_2O_5 . The amount (or 'loading') of DMT-HEG spacers on the surface of a fused silica fiber was ~1 nmol/fiber (48 mm in length).

Synthesis of oligonucleotides on optical fibers

Approximately 10 functionalized DMT-HEG-GOPS fibers (48 mm in length) were placed in a standard 10 μmol scale Applied Biosystems synthesis column and capped with acetic anhydride prior to DNA synthesis using the ABI supplied cycle. Detritylation was performed with 3% trichloroacetic acid in

dichloroethane. Activation of phosphoramidites was achieved with 0.5 M tetrazole in acetonitrile. Reagents for capping were as follows: Cap A, phenoxyacetyl anhydride Cap A reagent from Millipore (Mississauga, ON); and Cap B, 16% *N*-methylimidazole in THF (w/v). Iodine, 0.1 M, in THF/pyridine/water (25:20:2, v/v/v) was used for oxidations. Phenoxyacetyl protected dG, dC, dA phosphoramidite monomers were obtained from Millipore.

*N*⁶-phenoxyacetyl-3'-*O*-DMT-2'-deoxyadenosine-5'-*O*-[([beta]-cyanoethyl)*N,N*-diisopropyl]-phosphoramidite was prepared via standard protocols (39). The oligomers were deprotected with conc. NH₄OH solution for 2 h at room temperature. Following deprotection, the ammonia solution was collected, the column was washed with autoclaved water and the eluent was also kept. Quantitation of the eluents at 260 nm indicated that ~20% of the oligomers remained bound to the fiber surface.

Synthesis of branched oligonucleotides

The 'V' branched sequence **1** (Fig. 3) was synthesized on an Applied Biosystems 381A instrument using a 1 μmol scale synthesis cycle and [beta]-cyanoethylphosphoramidite chemistry (3,40). Purification, desalting, and analysis of the branched oligonucleotide **1** was accomplished by our detailed protocols (3,41). Typical isolated yields of this branched oligomer were 15-25% (~0.4-1.5 mg), as determined by absorption at 260 nm.

UV thermal denaturation and renaturation studies

Absorbance versus temperature profiles of the nucleic acid complexes (10 mM Tris, 50 mM MgCl₂, pH 7.3, 2.5 × 10⁻⁸ M ethidium bromide) were measured at 260 nm using a Varian Cary I UV-VIS spectrophotometer. Thermal denaturation profiles (i.e., melting curves) and thermal renaturation profiles (i.e., cooling curves) of each system of oligonucleotides were acquired at two temperature ramp rates, 0.5°C/min and 0.06°C/min. For each system of oligonucleotides, the denaturation and renaturation profiles provided identical results for the melting temperature (*T*_m) and showed no dependence on the temperature ramp rate used. Normalized plots were constructed according to the method of Kibler-Herzog *et al.* (42). All complexes showed sharp melting transitions. The values of *T*_m were determined from the first derivative of each thermal curve with an error in precision not greater than ±0.5°C based on variance in repeated experiments.

Figure 1. Schematic diagram of the apparatus used for fluorescence investigations of nucleic acid hybridization on the fiber optic sensor.

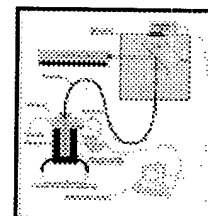


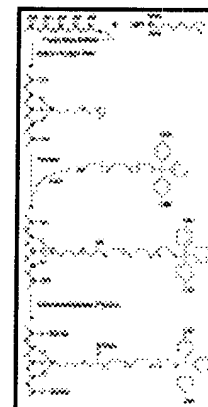
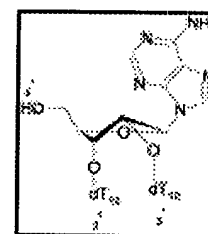
Figure 2. Derivatization of fused silica optical fibers.

Figure 3. The chemical structure of compound 1, a branched oligonucleotide with identical chains linked to the 2'- and 3'-positions of a ribose branch-point nucleoside, i.e., rA[(2'-5'-dT₁₀)/(3'-5'-dT₁₀)]. Ad, adenosine; Th, thymine. Two molecules of dT₁₀ hybridize with dA₁₀ to give the more common parallel (T*AT, Hoogsteen) triplex, whereas 1 forms a triplex in an antiparallel (T·AT, reversed-Hoogsteen) binding motif.



Instrument setup and fluorescent measurements

The instrument used for fluorescence intensity measurements was based on a fluorescence microscope as was previously described by Krull and co-workers (43) and shown in Figure 1. Radiation from an Ar⁺ laser operated at 488 nm was reflected by the dichroic mirror (495 nm cut-off) in the fluorescence microscope to a Zeiss 16* immersion lens with a numerical aperture of 0.5 (Empix Imaging, Mississauga, ON, Canada). The laser radiation exciting the immersion lens was coupled into a delivery fiber of similar numerical aperture (0.48) aligned beneath the objective. The light was totally internally reflected along the length of the delivery fiber to a sensing fiber functionalized with immobilized oligonucleotide. Coupling of the radiation between fibers was achieved by abutting the distal terminus of the delivery fiber to the proximal terminus of the sensing fiber. A loss in optical transmission of no more than 2% was observed for the coupled system. The termini of the teflon fiber coupler were designed as compression-fit ends which provided a solution-tight seal that prevented contaminants from diffusing into the fiber coupler and causing drift in the analytical signal. The sensing fiber was placed in a small volume, stop-flow, stainless steel hybridization chamber (1.5 mm i.d. * 48 mm) which provided a solution volume of 79 μ l immediately surrounding the sensing fiber. The temperature of the hybridization cell was controlled by placing the cell in a thermostated housing. The temperature of the solutions in the hybridization cell were accurately determined ($\pm 0.2^\circ\text{C}$) by use of a glass encapsulated thermistor incorporated into the hybridization cell. Solutions containing hybridization buffer, ethidium bromide, and complementary nucleic acid sequences were delivered to the hybridization cell and sensing fiber by use of a peristaltic pump. In all cases, a hybridization buffer/dye solution of 10 mM Tris, 50 mM MgCl₂, 2.5×10^{-8} M ethidium bromide at pH 7.3 was used unless otherwise specified. Fluorescence emission from ethidium bromide that was intercalated into immobilized nucleic acid complexes was totally internally reflected within the sensing fiber and directed towards a photomultiplier tube, where the fluorescence intensity could be quantitatively measured. Drift caused by variations in the efficiency of optical coupling, laser intensity and photomultiplier gain were obviated by normalization of all signals to that of a standard

solution of ethidium bromide at 25°C prior to and at the completion of each analysis.

PAGE mobility retardation assay

The solutions of oligonucleotides (10 μ l of 30% sucrose/50 mM MgCl_2) were incubated at 4°C (96 h) then loaded onto a non-denaturing 15% polyacrylamide gel (90 mM Tris-borate/50 mM MgCl_2 , pH 8.0). The gels were run at 12.5 mA for 12 h after which the bands in the gel were visualized and photographed by UV illumination followed by ethidium bromide staining.

RESULTS AND DISCUSSION

A goal of this research endeavor was to create a rapid and reliable assay for the detection of triple-helical nucleic acid formation as an extension of work initiated for the detection of duplex formation (14). As a starting point, we chose to investigate the parallel and antiparallel T/AT triplexes as these have been well documented in the literature. Branched nucleic acids as described by Damha *et al.* (3,40) were also used in this study as their unique architecture has been shown to stabilize reversed-Hoogsteen T·AT (antiparallel) triplexes (44). The advantage provided by our optical sensor technology over standard fluorometric work include the low detection limits, reusability and reliability of the device, the non-destructive nature of the assay (where samples may be collected and re-used) and this approach readily lends itself to automation, thereby negating the requirement of highly skilled technicians to carry out the assay.

Figure 4. Fluorescent intensity as a function of temperature dA₁₀ functionalized

sensors challenged with dT₁₀. Response of the optical sensor to 2.5×10^{-8} M ethidium bromide (solid star). Response of the optical sensor with 5' → 3'-fiber immobilized dA₁₀ to 40 pmol of linear dT₁₀ in the presence of 2.5×10^{-8} M ethidium bromide (closed circle). Response of the optical sensor with 3' → 5'-fiber immobilized dA₁₀ to 40 pmol of linear dT₁₀ in the presence of 2.5×10^{-8} M ethidium bromide (cross in open circle). Cooling profile of the same nucleic acid system in bulk solution by measurement of absorbance at 260 nm (thick broken line).



Immobilization of oligonucleotides onto optical fibers

The hydroxylated surfaces of the fused silica optical fibers were activated by reaction with GOPS followed by extension with a DMT-HEG linker (Fig. 2). This provides a derivatized surface consisting of a hydrophilic, long-chain spacer arm with a DMT-protected hydroxyl terminus onto which oligonucleotides may be assembled via solid-phase phosphoramidite synthesis (Materials and Methods). This linker was chosen because it is stable to standard oligonucleotide deprotection conditions (37), and provides a fluid environment which facilitates hybridization between immobilized DNA strands and the target strands in solution (47).

Parallel and anti-parallel T-AT triplex considerations

Formation of the intermolecular triplex 2·dT₁₀:dA₁₀ may be characterized by a variety of techniques including UV melting studies, molecular modeling, circular dichroism and NMR spectroscopy (48,49).

In the pyrimidine motif, the third dT₁₀ strand interacts by means of Hoogsteen hydrogen bonds with the dA₁₀ strand in target duplex, and is oriented parallel to it. In melting experiments (Mg²⁺ buffer), the triplex 2*dT₁₀:dA₁₀ has two resolved transitions, one for dissociation of the third strand from the duplex, i.e., dT₁₀*dA₁₀:dT₁₀ → dT₁₀ + dA₁₀:dT₁₀ ($T_m = 18^\circ\text{C}$), and one for dissociation of the duplex into its component strands, i.e., dA₁₀:dT₁₀ → dA₁₀ + dT₁₀ ($T_m = 32^\circ\text{C}$) (50). Thus association of the third (dT₁₀) strand with the duplex (dA₁₀:dT₁₀) is thermodynamically weaker than duplex formation itself (51).

Work done in our laboratories has shown that branched oligonucleotides are useful probes for stabilizing triplex DNA (44). The branched oligomer **1** (Fig. 3) for instance, binds to dA₁₀ via reversed-Hoogsteen interactions to give a three-stranded complex in which both dT₁₀ strands are antiparallel to the purine (dA₁₀) strand. The formation of this triplex was induced by linkage of two dT₁₀ strands through their 5'-ends via coupling to riboadenosine at the neighboring 2' and 3' oxygen atoms. Although this motif had been observed for T-AT bases in complexes dominated by pur-pur:py bonding (e.g., G-GC, A-AT) (52,53), it has only been observed recently for dT_n:dA_n complexes (44,54). Thermal denaturation and renaturation profiles of a mixture of **1** and dA₁₀ (1:1) in Mg²⁺ buffer show a single transition from bound to unbound complex (44), consistent with its formation involving a single bimolecular step, i.e., **1** + dA₁₀ → triplex **1**:dA₁₀ ($T_m = 35^\circ\text{C}$).

Figure 5. Fluorescent intensity as a function of temperature for the mixed base sequence icosanucleotide functionalized fibers. Upper curve: response of the optical sensor to 20 pmol of linear complement icosanucleotide in the presence of 2.5×10^{-8} M ethidium bromide. Lower curve: response of the optical sensor to 2.5×10^{-8} M ethidium bromide.

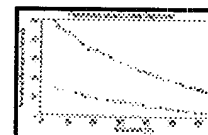
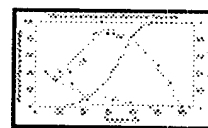


Figure 6. Fluorescent intensity as a function of temperature for **1** using reversed orientation 3'-dA₁₀-5'-fiber derivatized sensors. Response of the optical sensor to 40 pmol of **1** in the presence of 2.5×10^{-8} M ethidium bromide (closed circle) and to the 2.5×10^{-8} M ethidium bromide solution alone (solid star). Cooling profile of the same nucleic acid system in bulk solution by measurement of absorbance at 260 nm (broken line).

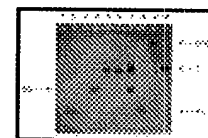


Triplex studies using derivatized optical fibers with normal (5'-dA₁₀-3'-fiber) oligonucleotide orientation

Decadeoxyadenylic acid (dA₁₀) was grown in the conventional 3' to 5' direction from the fiber surface. Solutions of hybridization buffer containing ethidium bromide, ethidium bromide with dT₁₀ or ethidium bromide with **1** were heated (~60°C) in the hybridization chamber containing the decaadenylic acid functionalized optical fibers and renaturation was followed spectroscopically. Fluorescence intensity as a function of temperature for 5'-dA₁₀-3'-fiber functionalized sensors challenged with dT₁₀/ethidium bromide is shown in Figure 4. As the temperature was lowered to 20°C, there was an increase in the fluorescence intensity due to the quantum yield enhancement of the duplex intercalated ethidium bromide.

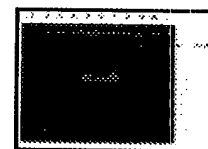
Upon further cooling, a decrease in the fluorescence intensity with decreasing temperature was observed, indicative of ligand exclusion due to triplex formation ($2 \cdot dT_{10} : dA_{10}$). In order to verify that triplex formation was alone responsible for the exclusion of the ethidium cation and the resulting decrease in fluorescence intensity, a control experiment was done using optical fibers functionalized with a 20 nt sequence of mixed base composition. Because this sequence lacked a pyrimidine (Py)_n or purine (Pu)_n stretch, only a double-stranded complex could form on the surface of the optical sensor upon binding to a complementary sequence. The hybridization experiment was carried out under the same conditions as for the dA_{10} functionalized fibers with the exception of the hybridization buffer (1 M NaCl, 50 mM PO_4^{2-} , pH 7.0). Intense fluorescence with a negative temperature coefficient was observed for the duplex system over the temperature range studied (10–65°C, $T_m = 73^\circ\text{C}$). The control experiment with ethidium bromide and no complementary oligonucleotide showed a negative temperature coefficient over the same temperature range with no such dramatic increase in intensity (Fig. 5).

Figure 7. Photograph of a UV-shadowed native polyacrylamide gel containing single strands, duplex and triple helical complexes of branched and linear controls. DNA samples were loaded in 50 mM $MgCl_2$, and 30% sucrose. Lane 4, dT_{10} ; lane 5, $dT_{10} : dA_{10}$ (1:1); lane 6, $dT_{10} : dA_{10}$ (2.5:1); lane 7, $dT_{10} : dA_{10}$ (4:1); lane 8, dA_{10} ; lane 9, **1** + dA_{10} ; lane 10, **1**. As can be noted, the $dT_{10} : dA_{10}$ triplex (lane 7) showed a slight retardation in the mobility relative to the corresponding duplex (lanes 5 and 6). The slowest mobility was observed in lane 9 for the branched triplex **1** : dA_{10} .



Interestingly, upon exposure of the optical sensor to the reversed-Hoogsteen forming **1**, no significant increase in fluorescence intensity over that of the ethidium bromide alone in solution was observed (data not shown). The geometrical constraints of compound **1** are such that, if a complex formed with the immobilized dA_{10} strand in this particular (fiber-3'- dA_{10} -5') orientation, the branch-point riboadenosine moiety would be oriented toward the fiber surface, and thus present a steric barrier to triplex formation. In order to facilitate the formation of the desired antiparallel branched triplex (and test whether steric interference surrounding the branch-point prevented triple-helical formation), an optical sensor having dA_{10} strands in the opposite orientation from the surface, i.e., fiber-5'- dA_{10} -3', was prepared.

Figure 8. Photograph of an ethidium bromide-stained native polyacrylamide gel (same gel as Fig. 7) containing single strands, duplex and triple helical complexes of branched and linear controls. DNA samples were loaded in 50 mM $MgCl_2$, and 30% sucrose. Lanes 4–10 are the same as those indicated in Figure 7. As can be noted, the $dT_{10} : dA_{10}$ triplex (lane 7) showed a slight retardation in the mobility relative to the corresponding duplex (lanes 5 and 6). The slowest mobility was observed in lane 9 for the branched triplex **1** : dA_{10} . Note that only the duplexes and triplexes showed ethidium bromide fluorescence.



Triplex studies using derivatized optical fibers with reversed (3'- dA_{10} -5'-fiber) oligonucleotide orientation

The fluorescence intensity versus temperature profile with dT_{10} shows an initial increase in fluorescence

intensity with decreasing temperature, indicative of duplex formation (Fig. 4). With further cooling of the system, the polarity of the fluorescence intensity temperature coefficient then inverts, indicative of triplex formation. Treatment of the optical sensor with **1** also provided a fluorescence intensity with a positive temperature coefficient at temperatures below the T_m (35°C), indicative of the formation of the reverse-Hoogsteen complex (Fig. 8).

The results of these experiments can be best understood by considering the two key competing factors which influence the net fluorescence intensity temperature coefficient. Firstly, the fluorescence quantum efficiency of the intercalant ligand bound to triple-stranded nucleic acids is greater than that of the ligand bound to double-stranded nucleic acid (36,45,46). This is the result of the triple-stranded structure being more rigid than the double-stranded nucleic acid structure, thereby providing superior shielding of the intercalated fluorophore from non-radiative collisional deactivation. In both cases, triplex and duplex, the quantum efficiency of the bound fluorophore increases with decreasing temperature (i.e., displays a negative temperature coefficient) owing to the overall reduction in the molecular motion in the system. The second factor influencing the net fluorescence emission is the binding efficiency of the intercalant ligand to each substrate type. Not as many ethidium cations can be accommodated per base triplet as per base pair. In addition, further exclusion of ethidium cation occurs with decreasing temperature in triple-helical nucleic acids, thereby providing a fluorescence intensity with a positive temperature coefficient. At low temperatures, the exclusion process dominates the fluorescence signal, thereby providing a means for elucidation of triple-strand formation.

In greater detail, it can be inferred from the data of Scaria and Shafer (36) that under these conditions of ionic strength and pH, a temperature below 25°C is required for the ethidium cation exclusion process to dominate the net fluorescence signal. Given that intercalation occurs at a maximum of every 2.8 base triplets and once per 2.4 base pairs at 25°C, a 14% reduction in the amount of intercalated ethidium occurs upon triple-strand formation. However, within the triplex structure, the fluorescence quantum yield of the remaining intercalated ethidium cation increases by 19% for the $S_1 \rightarrow S_0$ electronic transition, thereby resulting in a net fluorescence intensity change of +2.3%. Therefore, direct correlation between the T_m for triplex formation and the onset of fluorescence emission with a positive temperature coefficient will be observed for systems of nucleic acids which have T_m values at or below ~25°C. This is consistent with our findings (Fig. 4) whereby the decrease in fluorescence intensity from the sensor correlates well with the temperature at which dT_{10} associates to the dT_{10}/dA_{10} duplex ($T_m = 18^\circ\text{C}$).

Although the transition for triple-strand formation between **1** and the immobilized dA_{10} occurs at 35°C (Fig. 6), a decrease in fluorescence intensity was not observed until the system was cooled to below ~25°C. In this regard, our fluorescence studies involving ethidium bromide binding to triple-helices is in full agreement with several earlier findings. Our system is then limited in terms of being able to identify the duplex to triplex transition temperature for nucleic acid systems with T_m values at or below 25°C. This does not, however, limit the applicability of this technology in terms of being a useful strategy to identify triplex formation.

It is also interesting to note in Figure 6, for the binding of **1** with immobilized dA_{10} , a significant fluorescence intensity is observed over the temperature range from ~50 to 60°C. This is indicative of the presence of intercalated ethidium cation. This is contrary to data presented in the UV denaturation/renaturation profiles for the same oligonucleotide system in solution where no significant quantity of complex formation exists over that temperature regime. A possible explanation for this unusual observation is that the ionic strength at or near the surface of the optical sensor may be greater

than that of bulk solution owing to the presence of the immobilized polyanionic nucleic acid strands. As such, a shift in the T_m to higher temperatures would be expected. This is consistent with our previously reported data where binding of dA₂₀ to immobilized dT₂₀ was found to have a T_m value greater than that of the same oligonucleotide system in solution (14).

PAGE mobility retardation assay

Gel-shift experiments confirmed the interaction of ethidium bromide with the complexes observed in these studies. The electrophoretic mobility of the dT₁₀:dA₁₀ duplex, both the Hoogsteen and reverse-Hoogsteen paired T·AT triplexes, and that of their component strands, was studied at 4°C. Following electrophoresis, the gels were visualized by UV shadowing, and by staining with ethidium bromide (Figs 7 and 8, respectively). The Hoogsteen triplex migrated more slowly than the duplex while the reversed-Hoogsteen triplex showed the slowest mobility of all, which is characteristic of branched nucleic acid structures (55). Association of 1 and dA₁₀ was quantitative as evidenced by the complete disappearance of compound 1 and dA₁₀, when mixed in equimolar amounts, as visualized in the gel (Fig. 7). The stoichiometry of interaction between dT₁₀ and dA₁₀ for the duplex and Hoogsteen triplex was also confirmed by studies at different concentrations of the two oligonucleotides. When stained with ethidium bromide and illuminated by a UV lamp, fluorescence was observed only in the bands corresponding to the complexes, not single strands (Fig. 8). This is consistent with the well-known intercalative binding motif of ethidium bromide (56). As previously suggested by the biosensor studies, the 1/dA₁₀ reverse-Hoogsteen triplex gave the lowest fluorescence intensity, which could be caused by the limited availability of ethidium binding sites in this complex.

Conclusions

In conclusion, a novel method for the detection of triple-helical nucleic acid formation has been demonstrated. The complementary oligonucleotide, dT₁₀, which was expected to hybridize via a double-stranded Watson-Crick motif to immobilized dA₁₀ provided a fluorescence intensity with a negative temperature coefficient upon cooling the system below the duplex melting temperature ($T_m = 32^\circ\text{C}$). Upon further cooling, to the point where Hoogsteen T*AT triple-helix formation occurred, a fluorescence intensity change with a positive temperature coefficient was observed as a result of exclusion of the ligand from the triplex structure. Similar results were observed for triplex formation between dT₁₀ and the immobilized dA₁₀ sequence in both the normal (fiber-3' → 5') orientation and the reversed (fiber-5' → 3') orientation. The reversed-Hoogsteen T·AT triplex formed with 1 and the immobilized dA₁₀ grown in reversed orientation (fiber-5' → 3') also provided a fluorescence intensity with a positive temperature coefficient, consistent with triplex formation and ligand exclusion. Correlation between the triplex T_m and the temperature at which the temperature coefficient of the fluorescence intensity changes from negative to positive may be observed for nucleic acid systems with a triplex T_m below ~25°C. Determination of triplex formation may be done rapidly (in minutes) by setting the initial temperature of the system to that of the triplex T_m and then slowly cooling the system (-0.5°C/min) for a few minutes to determine the fluorescence intensity temperature coefficient.

Further studies will be directed to expanding the triple-helix sequence context, investigations of mismatch sensitivity, and developing less limiting fluorescent dyes. Optical sensors with covalently bound intercalant have been created in our laboratories which provide a reagentless sensing system with fast

response times (<6 min for full analytical response) for double-strand formation. Investigations of triplex formation on these reagentless sensors will also be evaluated in diagnostic assays, as they eliminate the problem of doubled-stranded DNA in the sample solution (e.g., in a biological sample) procuring all of the intercalant present in the buffer solution.

ACKNOWLEDGEMENTS

We acknowledge the Natural Science and Engineering Council of Canada (NSERC), and Le Fonds pour la Formation des Chercheurs et l'Aide à la Recherche (FCAR) for support of this work. We also acknowledge additional support from NSERC for a postgraduate scholarship awarded to P.A.E.P. and a postdoctoral scholarship awarded to R.H.E.H. We would also like to acknowledge Sebastien Robidoux for supplying the thermal renaturation data presented herein.

REFERENCES

- 1 Thomas, D. (1995) *BioTechnology*, **13**, 474-481.
- 2 Laval, J.M., Chopineau, J. and Thomas, D. (1995) *Trends Biotechnol.*, **13**, 474-481.
- 3 Hudson, R.H.E. and Damha, M.J. (1993) *J. Am. Chem. Soc.*, **115**, 2119-2124.
- 4 Horn, T. and Urdea, M.S. (1989) *Nucleic Acids Res.*, **17**, 6959-6967. [MEDLINE Abstract](#)
- 5 Kool, E.T. (1996) *Annu. Rev. Biophys. Biomolec. Struct.*, **25**, 1-28.
- 6 Chen, G. and Seeman, N.C. (1991) *Nature*, **350**, 631-633.
- 7 Ray, R.A., Cooper, P.J. and Wakefield, A.J. (1995) *BioTechnology*, **13**, 445-447.
- 8 McGown, L.B., Joseph, M.J., Pitner, J.B., Vonk, G.P. and Linn, C.P. (1995) *Anal. Chem.*, **67**, A663-A668.
- 9 Millan, K.M., Saraullo, A. and Mikkelsen, S.R. (1994) *Anal. Chem.*, **66**, 2943-2948. [MEDLINE Abstract](#)
- 10 Su, H.B., Chong, S. and Thompson, M. (1996) *Langmuir*, **12**, 2247-2255.
- 11 Nilsson, P., Persson, B., Uhlén, M. and Nygren, P. (1995) *Anal. Biochem.*, **22**, 400-408.
- 12 Bates, P.J., Dosanjih, H.S., Kumar, S., Jenkins, T.C., Laughton, C.A. and Neidle, S. (1995) *Nucleic Acids Res.*, **23**, 3627-3632. [MEDLINE Abstract](#)
- 13 Mirkin, C.A., Letsinger, R.L., Mucic, R.C. and Storhoff, J. (1996) *Nature*, **382**, 607-609. [MEDLINE Abstract](#)
- 14 Piunno, P.A.E., Krull, U.J., Hudson, R.H.E., Damha, M.J. and Cohen, H. (1995) *Anal. Chem.*, **67**, 2635-2643.
- 15 Graham, C.R., Leslie, D. and Squirrell, D.J. (1992) *Biosens. Bioelectron.*, **7**, 487-493. [MEDLINE Abstract](#)
- 16 Abel, A.P., Weller, M.G., Duveneck, G.L., Ehrat, M. and Widmer, H.M. (1996) *Anal. Chem.*, **68**,

2905-2912. [MEDLINE Abstract](#)

17 Strobel, S.A., Doucettstamm, L.A., Riba, L., Housman, D.E. and Dervan, P.B. (1991) *Science*, **254**, 1639-1642. [MEDLINE Abstract](#)

18 Maher, L.J., Dervan, P.B. and Wold, B. (1992) *Biochemistry*, **31**, 70-81. [MEDLINE Abstract](#)

19 Duvalvalentin, G., Thoung, N.T. and Hélène, C. (1992) *Proc. Nat. Acad. Sci. USA*, **89**, 504-508.

20 Stull, R.A. and Szoka, F.C., Jr (1995) *Pharmaceut. Res.*, **12**, 465-483.

21 Ito, T., Smith, C.L. and Cantor, C.R. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 495-498. [MEDLINE Abstract](#)

22 Wang, S.H., Friedman, A.E., Kool, E.T. (1995) *Biochemistry* **34**, 9774-9784.

23 Hoogsteen, K. (1959) *Acta Crystallogr.*, **12**, 822-823.

24 Felsenfeld, G. Davies, D.R. and Rich, A. (1957) *J. Am. Chem. Soc.*, **79**, 2023-2024.

25 Moser, H.E. and Dervan, P.B. (1987) *Science*, **238**, 645-650. [MEDLINE Abstract](#)

26 Praseuth, D., Perrouault, L., Doan, T.L., Chassignol, M., Thuong, N. and Hélène, C. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1349-1353. [MEDLINE Abstract](#)

27 Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J. and Hogan, M.E. (1988) *Science*, **241**, 456-459. [MEDLINE Abstract](#)

28 Beal, P.A. and Dervan, P.B. (1991) *Science*, **251**, 1360-1363. [MEDLINE Abstract](#)

29 Miller, J.H. and Sobell, H.M. (1966) *Proc. Natl. Acad. Sci. USA*, **55**, 1201-1205. [MEDLINE Abstract](#)

30 Morgan, A.R. and Wells, R.D. (1968) *J. Mol. Biol.*, **37**, 63-80. [MEDLINE Abstract](#)

31 Lee, J.S., Johnson, D.A. and Moran, A.R. (1979) *Nucleic Acids Res.*, **6**, 3073-3091. [MEDLINE Abstract](#)

32 Hoogsteen, K. (1963) *Acta Crystallogr.* **16**, 907-916.

33 Waring, M.J. (1966) *Biochem. J.*, **143**, 483-486.

34 LePecq, J.B. and Paoletti, C. (1967) *J. Mol. Biol.*, **27**, 87-106. [MEDLINE Abstract](#)

35 Mergny, J.L., Collier, D., Rougée, M., Montenay-Garestier, T. and Hélène, C. (1991) *Nucleic Acids Res.*, **19**, 1521-1526. [MEDLINE Abstract](#)

36 Scaria, P.V. and Shafer, R.H. (1991) *J. Biol. Chem.*, **266**, 5417-5423. [MEDLINE Abstract](#)

37 Maskos, U. and Southern, E.M. (1992) *Nucleic Acids Res.*, **20**, 1679-1684. [MEDLINE Abstract](#)

38 Pon, R.T. (1993) In Agrawal, S. (ed.) *Methods in Molecular Biology-Protocols for Oligonucleotides and Analogs*. Humana Press, Totowa, NJ, Vol 20, 465-496.

39 Van de Sande, J.H., Ramsing, N.B., Germann, M.W., Elhorst, W., Kalisch, B.W., Kitzing, E.V., Pon,

- R.T., Clegg, R.C. and Jovin, T.M. (1988) *Science* **241**, 551-557. [MEDLINE Abstract](#)
- 40 Damha, M.J., Ganeshan, K., Hudson, R.H.E. and Zabarylo, S.V. (1992) *Nucleic Acids Res.*, **20**, 6565-6573. [MEDLINE Abstract](#)
- 41 Damha, M.J. and Ogilvie, K.K., (1993) Agrawal, S. (ed.) *Methods in Molecular Biology-Protocols for Oligonucleotides and Analogs*. Humana Press, Totowa, NJ, Vol 20, pp 81-114.
- 42 Kibler-Herzog, L., Zon, G., Whittier, G., Shaikh, M. and Wilson, W.D. (1993) *Anti-Cancer Drug Des.*, **8**, 65-79.
- 43 Brennan, J.D., Brown, R.S., McClintock, C.P. and Krull, U.J. (1991) *Anal. Chim. Acta*, **237**, 253-263.
- 44 Hudson, R.H.E., Uddin, A. and Damha, M.J. (1995) *J. Am. Chem. Soc.*, **117**, 12470-12477.
- 45 Salunkhe, M., Wu, T. and Letsinger, R.L. (1992) *J. Am. Chem. Soc.*, **114**, 8768-8772.
- 46 Fang, Y., Bai, C.L., Zhang, P.C., Cao, E.H. and Tang, Y.Q. (1994) *Science In China (Ser. B)*, **37**, 1306-1312.
- 47 Wolf, S.F., Haines, L., Fisch, J., Kremsky, J.N., Dougherty, J.P. and Jacobs, K. (1987) *Nucleic Acids Res.*, **15**, 2911-2926. [MEDLINE Abstract](#)
- 48 Hopkins, H.P., Hamilton, D.D. Wilson, W.D. and Zon, G. (1993) *J. Phys. Chem.*, **24**, 6555-6563.
- 49 Radhakrishnan, I., Patel, D.J., Priestly, E.S., Nash, H.M. and Dervan P.B. (1993) *Biochemistry*, **32**, 11228-11234. [MEDLINE Abstract](#)
- 50 Pilch, D., Levenson, C., Shafer, R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1942-1946. [MEDLINE Abstract](#)
- 51 Pilch, D.S., Brousseau, R. and Shafer, R.H. (1990) *Nucleic Acids Res.*, **18**, 5743-5750. [MEDLINE Abstract](#)
- 52 Fox, K.R. (1994) *Nucleic Acids Res.*, **22**, 2016-2021. [MEDLINE Abstract](#)
- 53 Cassidy, S.A., Strekowski, L., Wilson, W.D., and Fox, K.R. (1994) *Biochemistry*, **33**, 15338-15347. [MEDLINE Abstract](#)
- 54 Dagneaux, C., Liquier, J. and Taillandier, E. (1995) *Biochemistry*, **34**, 14815-14818. [MEDLINE Abstract](#)
- 55 Wallace, J.C. and Edmons, M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 950-954. [MEDLINE Abstract](#)
- 56 Lim, C.S. (1994) *BioTechniques*, **17**, 626-628. [MEDLINE Abstract](#)
-

*To whom correspondence should be addressed. (M.J.D.) Tel: +1 514 398 7552; Fax: +1 514 398 3797; Email: damha@omc.lan.mcgill.ca (U.J.K.) Tel: +1 905 828 5437; Fax: +1 905 828 5425; E-mail: ukrull@credit.erin.utoronto.ca

[†]Present address: Department of Chemistry, University of Western Ontario, London, Ontario, Canada
The authors wish it to be known that, in their opinion, the first two authors should be considered as joint

First Authors

- [Abstract of this Article](#)
- [Reprint \(PDF\) Version of this Article](#)
- Similar articles found in:
 - [Nucl. Acids. Res. Online](#)
 - [PubMed](#)
- [PubMed Citation](#)
- Search Medline for articles by:
 - [Uddin, A. H.](#) || [Krull, U. J.](#)
- Alert me when:
 - [new articles cite this article](#)
- [Download to Citation Manager](#)

[HOME](#) [HELP](#) [FEEDBACK](#) [SUBSCRIPTIONS](#) [ARCHIVE](#) [SEARCH](#) [SEARCH RESULT](#)